CHAPTER III
RESEARCH METHODOLOGY

This chapter presents the research design and methodology. It begins with a description of the research site, followed by a detailed explanation of the research design, participants, data collection procedures, teaching program, including instructional design, and data analysis. The chapter concludes by addressing ethical considerations.

3.1 Research Site

This study was conducted at a private university in East Java, Indonesia, which was selected as the research site for several compelling reasons. First, the researcher has convenient access to the institution, having previously been both a student and an instructor there. This familiarity with the university environment and its participants increased the feasibility of the research project. Second, the study aligns with one of the university's core visions and missions: fostering proficient English writers capable of contributing to international networks. Thus, it facilitated the process of obtaining research approval. Third, the university's curriculum requires students at the tertiary level to compose argumentative texts, as highlighted by Dudley-Evans (2002), making the research context highly appropriate and relevant. Fourth, the potential participants would benefit from training to improve their English writing proficiency, as they are encouraged to prepare scholarly manuscripts for publication before graduation. Furthermore, personal communication with teaching staff and several students prior to the commencement of the study (20 May 2024) indicated that students experienced persistent difficulties in constructing coherent and well-developed argumentative texts, while the use of ChatGPT was perceived as particularly challenging due to the lack of clear institutional policies regulating its application. Lastly, the university context is particularly suitable for teaching discussion texts and incorporating ChatGPT into the learning process. This approach is anticipated to enhance students'

comprehension of the writing pedagogy they will implement as future teacher candidates.

3.2 Research Design

This study investigated the integration of ChatGPT into a Genre-Based Approach (GBA) for teaching argumentative writing in an Indonesian EFL teacher education program. Given its focus on pedagogical practices situated in real classroom contexts and its aim to explore the participants' meaning-making processes, a qualitative research design was employed. This approach allows for a nuanced exploration of participants' experiences, interpretations, and perceptions, elements that cannot be captured through experimental or positivist methods (Kingsley et al., 2010; Widodo, 2014). Qualitative inquiry also supports methodological flexibility, enabling the researcher to respond to emergent insights and contextual variables as the study progresses.

This inquiry is situated within a constructivist paradigm, which assumes a relativist ontology, a subjectivist epistemology, and a hermeneutic methodology (Lincoln et al., 2024). From a relativist stance, this paradigm posits that multiple realities exist, shaped by individuals' unique experiences and cultural contexts (Lincoln et al., 2024; Phakiti & Paltridge, 2015). Epistemologically, knowledge is co-constructed through interaction between the researcher and participants, who bring their own interpretations and values (Lincoln et al., 2024; Phakiti & Paltridge, 2015). Methodologically, a hermeneutic process is employed, where understanding is built through iterative interpretation, dialogue, and reflexivity (Lincoln et al., 2024; Phakiti & Paltridge, 2015). This paradigm aligns with the study's goal of exploring how students and teachers make sense of and interact with GenAI tools, specifically ChatGPT, within the genre-based writing classroom.

In line with this paradigm, a case study design was adopted to examine the phenomenon within its real-life setting (Yin, 2018). A case study enables an in-depth investigation of a bounded system, an instructional program involving sophomore EFL students at a public university in Indonesia. These students, majoring in English education, possess pre-intermediate levels of English

proficiency. The case study approach was selected not only to generate a comprehensive understanding of instructional practices and learning experiences but also to allow the researcher, who concurrently served as both teacher and instructional designer, to implement an intervention aligned with students' linguistic needs and pedagogical goals (Stake, 1995; Gast, 2009). This dual role enabled iterative adaptation of the instructional design in response to the evolving needs and responses of learners. Figure 3.1 displays the research procedure of the current study.

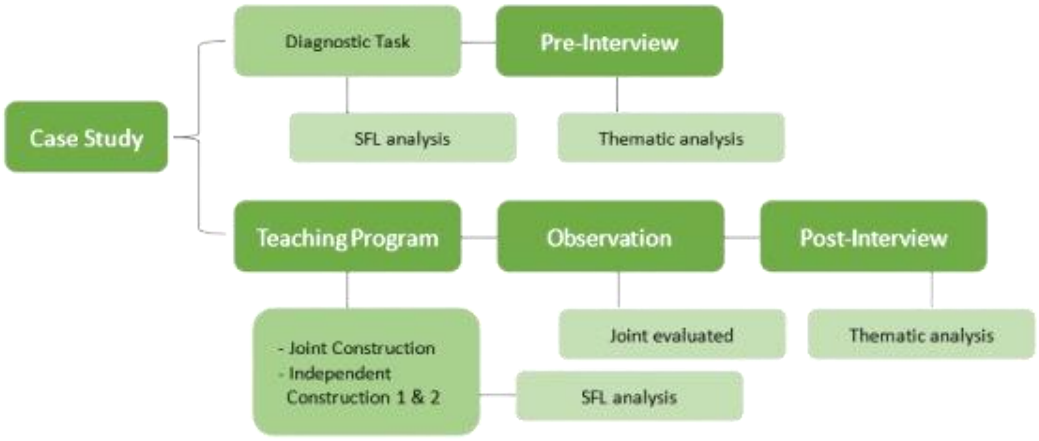

Figure 3.1. Research Procedure

Upon obtaining research approval and participants' consent, the study proceeded through the following stages: (1) administration of a diagnostic writing task in the form of a discussion text, followed by a pre-interview conducted the next day prior to the instructional intervention; (2) implementation of a two-cycle teaching program: Cycle 1 comprised four stages (Building the Context, Modelling/Deconstruction, Joint Construction, and Independent Construction), while Cycle 2 included two stages (Building the Context and Independent Construction). Classroom observations were conducted concurrently with the instructional sessions, assisted by a colleague; (3) students' writing data were collected from the two Independent Construction stages; and (4) a post-interview was conducted one day after the completion of Independent Construction in teaching cycle 2. Student writing samples were analyzed using Systemic Functional Linguistics (SFL), observation data were examined through collaborative

evaluation and discussion to compare and contrast findings, and interview data were analyzed thematically. These instruments facilitated triangulation and a rich portrayal of participants' developmental trajectories, perceptions, and challenges throughout the process.

Overall, this research design, anchored in the constructivist paradigm and operationalized through a case study approach, provided a robust framework for understanding the pedagogical integration of AI into genre-based writing instruction. By focusing on context, interaction, and the learners' evolving engagement with ChatGPT, the study aims to contribute to theory and practice in AI-assisted language education.

3.3 Participants

This study employed convenience sampling to recruit participants, capitalizing on the researcher's access to a cohort of students while prioritizing ethical research practices (Creswell & Guetterman, 2019). In a qualitative case study, a relatively small number of participants facilitates in-depth exploration of individual perspectives, ensures data richness, and supports the attainment of theoretical saturation while remaining manageable for detailed analysis (Casanave, 2015). Such an approach aligns with the nature of case study research, in which data are typically gathered over time, in considerable depth, and from a bounded number of participants and settings (Casanave, 2015). Accordingly, 15 sophomore students enrolled in an English education program at a private university in East Java, Indonesia, were recruited for the study. All participants had prior formal English instruction during their secondary education and were assessed at the pre-intermediate level based on their university entrance examination scores.

Ethical protocols were strictly observed throughout the recruitment and research process. Although the researcher had institutional access to the research site, informed consent remained a fundamental prerequisite. A detailed consent form was distributed to all prospective participants, clearly outlining the study's objectives, voluntary nature of participation, confidentiality safeguards, and the right to withdraw at any stage without penalty. Prior to the commencement of data

collection, the researcher provided a verbal explanation of the study to ensure participants' full understanding and informed agreement.

To verify participants' suitability for the intervention, a preliminary background interview was conducted. This step confirmed that none of the participants had previously received instruction that integrated ChatGPT or similar AI tools within a genre-based instruction. Following this, a diagnostic writing task was administered to assess their initial proficiency in composing argumentative texts. This diagnostic served as a baseline for evaluating their progress throughout the instructional intervention.

The main instructional program integrated ChatGPT into a genre-based approach (GBA) to teaching argumentative writing. After participating in stages of the Teaching and Learning Cycle, including joint construction, students completed independent writing tasks to demonstrate their individual development in argumentative text composition. These final texts served as post-intervention writing samples.

Table 3.1. Students' demographic background participated in the post-interviews

No.	Students	Gender	Scores		
			Diagnostic	Task I	Task II
1.	S1	Male	1,8	2,2	2,8
2.	S4	Male	1,2	2,3	2,7
3.	S6	Female	1,8	2,7	2,5
4.	S7	Female	1,5	2,7	2,7
5.	S10	Female	2,0	2,3	2,5
6.	S11	Male	2,3	2,8	2,8
7.	S13	Female	1,7	2,5	2,5

To gain deeper insights into the learners' experiences and writing processes, purposive sampling was employed to select participants for follow-up interviews (Creswell, 2014). As shown in Table 3.1, seven students were chosen based on their availability and classification as average-achieving writers, determined by their performance in the writing tasks. These interviews offered a more nuanced understanding of how the integration of ChatGPT influenced their learning, autonomy, and engagement with argumentative writing.

3.4 Data Collection Procedures

This study employed multiple data collection instruments to ensure comprehensive and triangulated insights into the integration of ChatGPT within a Genre-Based Approach to teaching argumentative writing. These instruments included a diagnostic writing task, two independent writing tasks, classroom observations, and semi-structured interviews. Each instrument was strategically selected to capture different dimensions of the participants' writing development, classroom engagement, and perceptions throughout the instructional intervention. Detailed descriptions of the implementation and procedures for each instrument are presented in the subsequent sections and in the next chapter. In adherence to ethical research practices, informed consent was obtained from all participants prior to data collection, ensuring voluntary participation, confidentiality, and the right to withdraw at any stage.

3.4.1 Diagnostic and independent writing tasks

Diagnostic assessment serves as an essential tool to identify students' current proficiency, strengths, and areas requiring improvement in relation to targeted learning outcomes (Lee & Sawaki, 2009). It plays a pivotal role in guiding instructional design by offering both macro- and micro-level insights into learners' needs. As reported in previous studies (e.g., Feez & Joyce, 1998; Lee, 2015), such assessments facilitate student learning by informing the planning of pedagogical strategies that are responsive to individual and collective learning profiles. Consistent with these principles, the diagnostic phase in this study allowed the researcher to determine the overall areas of development required by the class as a whole, as well as the specific needs of individual learners.

Prior to the implementation of the genre-based instructional intervention, participants completed a diagnostic writing task in which they were asked to compose an argumentative (discussion) text in a single classroom session. The writing prompt, collaboratively negotiated with the participants, was titled "Does social media do more harm than good?". This initial composition served as a baseline measure of students' argumentative writing abilities. The texts produced

during this session were collected and later subjected to detailed analysis, with selected samples included in the findings section to illustrate student performance prior to the intervention.

In addition to the diagnostic task, two independent writing tasks were embedded within the instructional cycles. The first was administered at the end of the first teaching cycle, while the second was integrated into the last stage of the second cycle. These tasks were designed to evaluate students' individual development in writing argumentative texts following scaffolded instruction, including interaction with ChatGPT. The data from these two tasks, along with the diagnostic task, constituted the primary sources of student writing used for analysis in this study. Although some texts from the joint construction stage were also analyzed and presented in the findings section.

Each writing task was conducted under controlled conditions in the classroom. Students were provided with a printed writing prompt (i.e., "*ChatGPT: A Friend or a Foe?*", the topic for independent writing task I, and "*Should we work or continue to study?*" for independent writing task II) and lined answer sheets. They were instructed not to collaborate with peers or access any digital resources during the writing session. For the two post-instruction tasks, participants were given access to interact with ChatGPT for feedback generation after they finished composing their drafts. This interaction was under supervision and consultation of the teacher (see Chapter IV for further detail).

3.4.2 Classroom observations

Classroom observation enables researchers to examine teacher-student interactions within various group configurations (e.g., whole class, pairs), the nature of the activities conducted, and the levels of student engagement (Spada, 2019). It provides a means to systematically document activities in social environments across diverse forms and contexts, thereby enriching and complementing other data sources (Simpson & Tuson, 2003). However, given that qualitative research often incorporates subjectivity to understand participants within specific settings, a method referred to as participant observation (Starfield, 2015), it is beneficial to

integrate peer observation as a strategy to triangulate and validate the researcher's own observations (Burns, 2015).

In this study, classroom observations were carried out by the researcher and a colleague. To ensure the smooth execution of their responsibilities, an observation protocol was designed in advance. Observations were conducted during every session, with the observers systematically recording all activities undertaken by both the researcher and the students. The primary focus was to document the researcher's actions and their impact on student behavior. Furthermore, the observation notes offered a comprehensive account of student engagement and behavior throughout each phase of the teaching process. These notes were documented in Google Docs for ease of storage and data confidentiality.

Table 3.2 The observation process

Classroom Interaction	Teacher-Students' Activities
Interaction types 1. Teacher-student 2. Teacher-all students 3. Teacher-small group 4. Student-student 5. Students in groups	1. What does the teacher say to the students? 2. What do students say to respond to the teacher? 3. What does the teacher say to respond to the students? 4. What do students do to respond to the teacher?

During the observation process, the researcher and their colleague documented all occurrences, with particular attention to two key aspects: (1) the classroom environment and (2) students' activities. Table 3.2 provides a summary of the observation methodology. For the classroom environment, the observers focused on elements such as the classroom layout, types of interactions, and emotional learning. In contrast, students' activities were recorded by noting their verbal expressions and actions during the teaching program.

3.4.3 Interviews

Interview as a research instrument or data collection perspective is a "data mining" (Brinkmann & Kvale, 2015, p. 57) or "harvesting psychologically and linguistically interesting responses" (Potter, 2004, p. 206). From this perspective,

the interview supports the researcher's goal of collecting data as the "interviewee can provide access to their internal or psychological worlds and lived experiences" (Prior, 2018, p. 229). As one of the interview methodologies, a semi-structured interview enables the researcher to deviate from the predefined interview protocol, allowing for a more in-depth exploration of participants' responses (Vandergrift, 2015). For this reason, semi-structured interviews were employed and conducted after the teaching program. Through interviews, students can clarify what knowledge they have gained in the program, what they believe has changed, and which elements of the instruction are in charge of the developments.

Within the framework of Yin's (2018) case study methodology, the semi-structured interview questions were devised following the guidelines established by Jacob and Furgerson (2012) for developing interview protocols and conducting interviews. The interview protocol, which can be found in the appendix, offers detailed guidance to the interviewer to ensure that essential information is not missed during the potentially stressful interview process (Creswell & Creswell, 2023). The interview questions were also inspired by ideas from a review paper by Yang and Kyun (2022). Consequently, these inquiries were congruent with the research objectives. The interviews were conducted in Bahasa Indonesia and audio-recorded with the participants' consent.

To gain a deeper understanding of students' involvement, writing processes, and the development of their genre knowledge and writing skills during the program, the researcher selected participants based on specific criteria. This data was used to triangulate findings from other sources. The interviews were conducted after the teaching program and involved students chosen based on two main criteria: their performance levels (low, moderate, and high scores, see the scoring framework in section 3.6.1 and the interviewees' demographic background in section 5.2 in Chapter V) and their availability for the interview. Students were informed that each interview would last no longer than 30 minutes.

3.5 Teaching Program

The teaching program was implemented at the university in the English department, where the participants were enrolled in the Academic Writing I course, which spanned approximately one semester. The teaching program was structured to include a diagnostic test during the first session and 9 sessions dedicated to two instructional cycles. In total, the program comprised 10 meetings, each with a maximum duration of one and a half hours. Table 3.3 provides an overview of the teaching program. The overview of the two instructional cycles is provided below, and further details will be discussed in Chapter IV.

Table 3.3 The Teaching Program Overview

Sessions	Activity	Description
1	Diagnostic test	Writing an argumentative text (i.e., discussion text) on a common topic.
	Teaching cycle 1	
2 & 3	Stage 1 Developing the context	1. Familiarizing the topic through several activities enhanced with ChatGPT, including reading, watching a video, vocabulary building, and group discussion. 2. Engaging students in an AI literacy framework a. Understanding ChatGPT b. Accessing ChatGPT c. Navigating information based on the topic d. Prompting on ChatGPT e. Corroborating generated information (including paraphrasing practices) 3. Reflective practices
4 & 5	Stage 2 Modelling & Deconstruction	1. Engaging students in several activities to develop their knowledge and skills required for discussion texts 2. Reflective practices
6 & 7	Stage 3 Joint construction	1. Planning, composing, revising and editing (after engaging in collaborative feedback practices: peer, teacher, ChatGPT) 2. Reflective practices
8	Stage 4 Independent construction	1. Writing a discussion text with a similar topic 2. Engaging in ChatGPT feedback practice under the teacher's supervision 3. Reflective practices

Sessions	Activity	Description
9	***Teaching cycle 2*** Stage 1 Developing the context	1. Familiarizing the topic through several activities enhanced with ChatGPT, including reading and language awareness enhancement 2. Reflective practices
10	Stage 2 Independent construction	1. Writing a discussion text with a general topic 2. Engaging in ChatGPT feedback practice under the teacher's supervision 3. Reflective practices

3.5.1 Teaching Cycles

The teaching program integrated the genre-based instruction with an AI technology, ChatGPT, for nine writing instruction sessions. Two key frameworks underpinned the instructional design: the Genre-based Approach (Rothery, 1994) and the AI Literacy Framework (Tseng & Warschauer, 2023). Thus, the instructional design was enhanced with appropriate ChatGPT uses as an additional tool for helping students develop their argumentative writing skills. Table 3.4 presents such an enhancement.

Table 3.4. Integration of ChatGPT into the Genre-based Approach

GBA stage	Instructional focus	AI use	AI literacy element
Developing the context	Introducing the topic and brainstorming	For vocabulary building, summary, and additional use.	Understanding Accessing Prompting Corroborating Incorporating
Modelling and deconstruction	Analyzing the model texts (including purpose, structure, and language features)	For text analysis (schematic structure and language features)	Understanding Accessing Prompting Corroborating Incorporating
Joint construction	Collaborative writing and feedback practices	For feedback and additional use	Understanding Accessing Prompting Corroborating Incorporating
Independent construction	Individual essay writing and feedback practices	For feedback and additional use	Understanding Accessing Prompting Corroborating Incorporating
Additional pedagogical stage: Reflection	Evaluating AI use and the writing process	-	-

The teaching/learning cycle comprises 2 cycles, adapted from Rothery (1994) and Emilia (2005). Cycle 1 includes building the context (the AI literacy framework is introduced in this stage), modelling and deconstruction, joint construction, and independent construction. Cycle 2 comprises only two stages: building the context and independent construction. The following is the description of each stage in the teaching/learning cycle. It is important to note that this description is only an overview; the detailed explanation will be discussed further in the next chapter, the enactment of instructional design.

3.5.1.1 Stage 1: Developing the context

In this stage, the teacher elicits what the learners think and know about the topic of discussion through several reading activities. This stage aims to familiarize students with the topic. In other words, this stage develops students' knowledge about the topic they will write about. Emilia (2016) argues that students will be ready to write when they know what to write. She further argues that in this stage, students develop their knowledge about the topic they are about to write, not to discuss the kind of text they will write (Emilia, 2016). In addition to developing the context for students, the teacher also introduces the AI literacy framework, equipping students to use ChatGPT in appropriate ways.

The AI Literacy Framework encompasses a set of core competencies essential for the effective use of AI technologies such as ChatGPT, including understanding, accessing, prompting, corroborating, and incorporating (Tseng & Warschauer, 2023). Following this framework, students in this study were systematically guided to develop a critical awareness of ChatGPT's capabilities and limitations (*understanding*), enabling them to assess the appropriateness of utilizing AI-generated content in specific contexts. Instruction emphasized purposeful engagement with ChatGPT aligned with instructional goals, particularly in the context of writing discussion texts (*accessing*). Students were trained to formulate and refine prompts iteratively to generate output that was relevant and aligned with their learning objectives (*prompting*).

Additionally, the instructional process included regular opportunities for students to evaluate the accuracy and reliability of AI-generated content, with explicit guidance on recognizing and addressing potential biases and misinformation (*corroborating*). Ethical use of AI was also a central component of the instruction; students were taught how to incorporate AI-generated content responsibly, including appropriate paraphrasing, citation practices, and acknowledgment of AI as a source (*incorporating*). Through this integrative approach, students were equipped with both the technical and ethical competencies necessary for meaningful engagement with AI in academic writing. Finally, students are engaged in reflective practice concerning their experience during this stage.

3.5.1.2 Stage 2: Modelling and deconstruction

In the modelling and deconstruction stage of the GBA employed in this study, students were supported in analyzing the structure and language of discussion texts through a series of three scaffolded activities enhanced with ChatGPT: (1) identifying key features of the genre, including its schematic structure and linguistic elements; (2) examining model texts to reinforce their understanding; and (3) engaging in guided practice with supplementary texts to consolidate their analytical skills. The scaffolded activity was adapted from Martin and Dreyfus' (2015) interactional design. Finally, students are engaged in reflective practice concerning their experience during this stage.

3.5.1.3 Stage 3: Joint construction

During the *joint construction* stage, teachers and students collaborate to create a text within the same genre. Adopting interactional design (Martin & Dreyfus, 2015), students are invited to actively engage in the writing process, which encompasses three key phases: planning, composing, and revising. During the planning phase, students gathered and organized relevant information to support their arguments. In the composing phase, they collaboratively developed well-structured texts with scaffolding and guidance from the teacher. The revising phase involved editing drafts based on feedback from peers, the teacher, and ChatGPT.

Notably, ChatGPT was utilized in this phase to provide additional feedback and assist with language refinement, thereby supporting students in enhancing the clarity, coherence, and accuracy of their written texts. At the end of this stage, students reflect on their experience.

3.5.1.4 Stage 4: Independent construction

In the final stage of the instructional cycle, students undertook the independent construction of a discussion text, drawing upon their experience cultivated during the preceding sessions (Martin & Dreyfus, 2015). This stage was strategically designed to evaluate students' capacity to independently produce a coherent and well-structured discussion text, with limited external assistance. In addition, the students employ ChatGPT to provide feedback on their writing draft and edit the draft under the consultation with the teacher. Finally, students are engaged in reflective practice concerning their experience during this stage.

As previously explained, Cycle 2 consists of two stages, Developing the Context and Independent Construction, both of which are similar to those implemented in Cycle 1.

3.6 Data Analysis

This section outlines the procedures for data analysis, which involved examining students' written texts, observing classroom activities, and analyzing students' interview responses. Recognizing the inherent subjectivity of qualitative research (Holliday, 2015), the study employed multiple strategies to ensure credibility and manage the etic perspective. Triangulation of data sources, respondent validation, and the careful formulation of evidence-based claims were implemented to manage potential bias and strengthen the trustworthiness of the findings.

Triangulation was achieved by collecting data from multiple sources, such as students' written texts, classroom observations, and interviews, to ensure a more comprehensive understanding (Hyland & Hyland, 2006). Respondent validation, also known as member checking (Lincoln & Guba, 1985), involved sharing data

with participants and inviting their feedback to confirm the authenticity of the interpretations and minimize potential researcher bias. Lastly, the formulation of appropriate claims reflects the study's aim not to establish definitive proof but to offer insights that may prompt critical reflection and a deeper understanding of the phenomena under investigation (Holliday, 2015).

However, several limitations of this study should also be acknowledged. First, as the teacher-researcher, there is potential for bias in instructional delivery and data interpretation, despite efforts to triangulate findings across multiple sources. Second, the relatively short duration of the intervention may limit the observation of sustained impacts on students' writing development. Third, reliance on self-reported data, including student reflections and interviews, may introduce social desirability or recall bias. These limitations highlight opportunities for future research, such as extending the intervention period, involving independent observers, and incorporating more objective measures of writing development to further validate and extend the findings of this study.

3.6.1 Scoring the written texts

In this study, scoring students' written texts was adapted from Rose's (2007, as cited in Emilia, 2016), Matthiessen et al. (1992), and Nagao's (2019) scoring rubric. As seen in Table 3.5, students' written texts were scored based on the following categories: (1) purpose, (2) staging (issue, argument for/against, conclusion), (3) support (supporting the claim), (4) field, (5) tenor, and (6) mode. For the last three categories, the scoring rubric is broken down into further detail, as can be seen in Table 3.6. Emilia (2016) argues that using the rubric provides detailed scoring and helps the assessor differentiate able students and those who cannot write well yet.

Table 3.5 Scoring rubric

Categories	**Scores**		
	Poor (=1)	Moderate (=2)	Good (=3)
Does the text present different opinions or perspectives on a specific issue? (Purpose)			

Categories	Scores		
	Poor (=1)	Moderate (=2)	Good (=3)
Does the text go through the appropriate stages (i.e., Issue, Argument for, Argument against, Conclusion? (Staging)			
Is the argument sufficiently supported by relevant information? (Support)			
Does the text make appropriate ideational meaning choices? (Field)			
Does the text make appropriate interpersonal meaning choices? (Tenor)			
Does the text make appropriate textual meaning choices? (Mode)			

Note: Adapted from Rose (2007, as cited in Emilia, 2016), Matthiessen et al. (1992), and Nagao (2019)

Table 3.6 Field, Tenor, and Mode: Breakdown categories

Breakdown Categories	Scores		
	Poor (=1)	Moderate (=2)	Good (=3)
Field			
Using various types of Processes(e.g., action, sensing, saying, relating) expressed through verb groups (including tense)			
Using various types of Participants (e.g., everyday, abstract, technical, nominalized, specific) expressed through e.g., noun groups			
Using various types of Circumstances (e.g., time, place, manner, reason) expressed through adverbials (e.g., adverb groups, prepositional phrases)			
Using conjunctions (coordinating and subordinating)			
Tenor			
Expressing attitudes (e.g., sharing feelings, appreciating qualities, making moral judgments)			
Engaging with the reader, with the discourse community, with other possibilities, with alternative perspectives, with 'layers of meaning' (through resources such as modality, citation practices, metaphor, simile, personification)			
Adjusting the strength (e.g., attitude, probability)			
Mode			
Using text connectives e.g., to sequence ideas (firstly, finally), to express causality (therefore, as a result), to clarify (for example, that is), to express condition (if, in that case), and to express concession and contrast (however, though, alternatively)			
Using correct spelling and punctuation to assist meaning in written texts			

Note: Adapted from Rose (2007, as cited in Emilia, 2016), Matthiessen et al. (1992), and Nagao (2019)

This scoring rubric is based on Systemic Functional Linguistics (Halliday, 1985). Systemic Functional Linguistics (SFL) is a language theory that describes language as a tool for meaning-making and deals with the text that conveys meaning and the relationships between texts and contexts (Halliday & Martin, 1993; Lock & Jones, 2011), making it as appliable linguistics (Martin et al., 2022; Halliday, 2008). As To et al. (2015) argued, SFL offers a strong framework for qualitative text analysis by providing theoretical insights and practical tools to interpret both meaning and authorial intent. It considers contextual factors and reveals how linguistic choices, such as word selection, syntax, theme, and lexical density, serve specific communicative purposes. Thus, SFL is adopted in this study as the analytical tool to analyse students' written texts to see the development of their argumentative writing skills, as it provides the teacher with a process to understand, interpret, and evaluate the students' texts.

Accordingly, the following framework was proposed as a guide in scoring students' written texts. As seen in Table 3.6, this scoring framework, grounded in SFL, evaluates students' discussion texts across six categories: purpose, staging, support, field, tenor, and mode. A score of 1 (poor) indicates limited performance, such as the failure to present multiple perspectives, lack of structural organization, insufficient support, and weak language choices. A moderate score (2) reflects partial achievement, where students show some awareness of structure and language use but with inconsistencies or limited depth. A score of 3 (good) demonstrates a well-structured text with a clear argumentative purpose, balanced perspectives, relevant support, and effective use of linguistic resources. This includes varied processes and participants (field), appropriate engagement and stance (tenor), and coherent textual organization with accurate mechanics (mode). The framework helps identify students' strengths and areas for development in producing effective argumentative writing.

Table 3.7. Students' writing performance framework based on SFL-informed rubric

Category	Score 1 (Poor)	Score 2 (Moderate)	Score 3 (Good)
Purpose	Fails to present differing perspectives or opinions on the issue.	Presents limited or unbalanced perspectives.	Clearly presents multiple, balanced perspectives on the issue.
Staging	Lacks expected structure (e.g., missing issue, arguments, or conclusion).	Includes some stages, but the sequence or clarity is weak.	Follows appropriate structure: clear issue, balanced arguments, and logical conclusion.
Support	Lacks supporting evidence or gives irrelevant examples.	Provides some relevant support but lacks depth.	Offers sufficient, relevant, and well-integrated supporting evidence.
Field	Limited use of process types, vague or concrete participants, minimal use of circumstances, and few or incorrect conjunctions.	Uses some variation of processes, participants, and circumstances; some conjunctions are used correctly.	Effectively uses a variety of processes, participants (including abstract/technical), and circumstances, with correct and strategic use of conjunctions.
Tenor	Minimal expression of attitudes, lacks engagement or stance, inappropriate use of modality or evaluative language.	Expresses limited attitude or stance; some engagement present but inconsistent.	Expresses clear attitudes, effectively engages with the audience and perspectives using appropriate modality and evaluative language.
Mode	Ideas poorly connected; incorrect or inconsistent use of connectives; frequent errors in spelling and punctuation that hinder meaning.	Some idea connections made; basic use of connectives; minor issues in spelling/punctuation.	Ideas well connected with appropriate text connectives; accurate spelling and punctuation enhance clarity.

Note: Adapted from Rose (2007, as cited in Emilia, 2016), Matthiessen et al. (1992), and Nagao (2019)

3.6.2 Analyzing the observation data

The observation data were compared between the data from the researcher and the colleague. The data is then recursively read for data triangulation. It is done to make a contrast and comparison of all data obtained from different sources and aims to enhance the validity of the conclusion of the study (Freebody, 2003). As the observation involves two observers (the teacher and his colleague), the data are subjected to an interrater agreement process. To ensure the trustworthiness of the data, both parties engaged in a systematic discussion to compare and evaluate the data, identifying points of similarity and divergence until a consensus was achieved.

This collaborative process was intended to minimize potential biases and reduce the subjectivity of data recording and interpretation.

3.6.3 Analyzing the interview data

The interview data were initially transcribed under Widodo's (2014) methodological guidelines to ensure a meticulous and standardized approach to documenting verbal responses. The transcribed interviews were then subjected to thematic analysis. Following the framework proposed by Palys and Atchison (2014), initial codes were developed by identifying overarching themes by thoroughly reading the data. Subsequently, the data underwent an iterative, detailed examination to identify specific themes relevant to the research questions. These themes were carefully defined and labelled to enhance data organization and facilitate the analytical process. The interview data analysis findings were subsequently shared with the interviewees to confirm that the results accurately reflect their experiences.

3.7 Ethical Considerations

Before commencing the study, ethical clearance was obtained from the Institutional Review Board (IRB) at Universitas Pendidikan Indonesia, as the research involved human participants. Following this approval, consent forms were distributed to prospective participants to ensure their voluntary participation. These documents were provided in Bahasa Indonesia to guarantee clarity and full comprehension of the study's purpose, procedures, and potential benefits. Participants were clearly informed of their right to decline or withdraw from the study at any point without facing any unintended negative consequences, in accordance with ethical research standards (Beach & Eriksson, 2010).