GENDER DIFFERENCES IN STUDENTS' CRITICAL THINKING IN ARGUMENTATIVE WRITING

THESIS

In Partial Fulfillment of the Requirement for Master's
Degree in English Language Education

By:

RINALDY ALIDIN

202210560211032

DEPARTMENT OF ENGLISH LANGUAGE EDUCATION

THE DIRECTORATE OF GRADUATE PROGRAM

UNIVERSITAS MUHAMMADIYAH MALANG

2024

GENDER DIFFERENCES IN STUDENTS' CRITICAL THINKING IN ARGUMENTATIVE WRITING

by

RINALDY ALIDIN
202210560211032

Accepted on

Saturday, 13th July 2024

Advisor I	Advisor II
Dr. Estu Widodo	**Prof. Dwi Poedjiastutie, Ph.D**

Director of the Graduate Program	Head of Department
Prof. Latipun, Ph.D	**Dr. Estu Widodo**

T H E S I S

Written by:

RINALDY ALIDIN
202210560211032

Defended in front of the examiners
on **Saturday, 13th July 2024** and it was
decided that it fulfilled the requirements to get
the master's degree in English Language Education
at the Graduate Program of Universitas Muhammadiyah Malang

The Examiners

Chief	:	**Dr. Estu Widodo**
Secretary	:	**Prof. Dwi Poedjiastutie, Ph.D**
1st Examiner	:	**Assoc. Prof. Dr. Sudiran**
2nd Examiner	:	**Dr. Rina Wahyu Setyaningrum**

LETTER OF STATEMENT

I, the undersigned:

Name : **RINALDY ALIDIN**

NIM : **202210560211032**

Study Program : Masters in English Education

Hereby, declare that :

1. The thesis entitled: **GENDER DIFFERENCES IN STUDENTS' CRITICAL THINKING IN ARGUMENTATIVE WRITING** is my original work and contains no one's scientific paper that may be proposed to achieve an academic degree at any universities. Beside, there is no other's idea or citation except those which have been quoted and mentioned at the bibliography.
2. If this thesis is proven as a form of **PLAGIALISM** in this thesis, I am willing to accept the consequences including accepting the **CANCELLATION OF THE GRANTING OF MASTER DEGREE** and undergoing any Procedures required by the prevailing law.
3. This thesis can be used for literature review which can be accessed by others freely **NON-EXCLUSIVE ROYALTY**.

Thus, this statement is made truthfully to be used as appropriate.

Malang, 13th July 2024

The Writer,

RINALDY ALIDIN

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

GENDER DIFFERENCES IN STUDENTS' CRITICAL THINKING IN ARGUMENTATIVE WRITING

Rinaldy Alidin

ralidin@webmail.umm.ac.id

Department of English Language Education, Universitas Muhammadiyah Malang, East Java, Indonesia

Abstract

Critical thinking is an essential skill in the 21st century. This study aimed to measure the differences between male and female students in critical thinking through argumentative writing. In this study, the researcher synthesized the instrument score based on the theories of Facione (2015) and Toulmin et al. (2002). Then, this study used two raters to score the students' argumentative essays. This study's design employed an ex post facto approach with a quantitative descriptive method, involving 30 student essays. However, the study employed the quota sampling technique, which consisted of 15 essays written by male students and 15 essays written by female students. The results and findings in this study showed a high level of agreement between two raters in scoring six indicators of critical thinking in the inter-rater reliability test. The differences between male and female students in descriptive statistics revealed that male students outperformed female students in four indicators: interpretation, analysis, inference, and self-regulation. However, female students outperformed male students in two indicators: explanation and evaluation. The researcher used the non-parametric Mann-Whitney test in SPSS to analyze the results from the student scores. Then, the results showed that there was no significant difference between male and female students in critical thinking.

Keywords: Argumentative essay; Critical thinking; Gender differences.

GENDER DIFFERENCES IN STUDENTS' CRITICAL THINKING IN ARGUMENTATIVE WRITING

Rinaldy Alidin

ralidin@webmail.umm.ac.id

Department of English Language Education, Universitas Muhammadiyah Malang, East Java, Indonesia

Abstract

Berpikir kritis adalah keterampilan penting di abad ke-21. Penelitian ini bertujuan untuk mengukur perbedaan antara siswa laki-laki dan perempuan dalam berpikir kritis melalui penulisan argumentatif. Dalam penelitian ini, peneliti mensintesis skor instrumen berdasarkan teori Facione (2015) dan Toulmin et al. (2002). Kemudian, penelitian ini menggunakan dua penilai untuk menilai esai argumentatif siswa. Desain penelitian ini menggunakan pendekatan ex post facto dengan metode deskriptif kuantitatif, yang melibatkan 30 esai siswa. Namun penelitian ini menggunakan teknik kuota sampling, yang terdiri dari 15 esai yang ditulis oleh siswa laki-laki dan 15 esai yang ditulis oleh siswa perempuan. Hasil dan temuan dalam penelitian ini menunjukkan tingkat kesepakatan yang tinggi antara dua penilai dalam menilai enam indikator berpikir kritis dalam uji reliabilitas antar penilai. Perbedaan antara siswa laki-laki dan perempuan dalam statistik deskriptif menunjukkan bahwa siswa laki-laki mengungguli siswa perempuan dalam empat indikator: interpretasi, analisis, inferensi, dan pengaturan diri. Namun siswa perempuan mengungguli siswa laki-laki dalam dua indikator yaitu penjelasan dan evaluasi. Peneliti menggunakan uji non-parametrik Mann-Whitney di SPSS untuk menganalisis hasil nilai siswa. Kemudian, hasil penelitian menunjukkan bahwa tidak terdapat perbedaan yang signifikan antara siswa laki-laki dan perempuan dalam berpikir kritis.

Keywords: Tulisan argumentatif; Berfikir kritis; Perbedaaan gender.

CHAPTER I

INTRODUCTION

This chapter discusses six main points. They are the background of the study, the research question, the objective of the study, the significance of the study, the scope and limitation and the definition of the key terms.

1.1 The Background of The Study

Critical thinking is one of the essential skills needed in the 21st century, mainly in language education. It is a component of a larger sustainability discourse in which education is key (UNESCO, 2017). It occurs when someone uses logic and a reflective point of view to try to lessen skepticism about a certain topic in a specific context. Critical thinking is the capacity to apply higher cognitive skills such as analysis, synthesis, self-reflection, perspective-taking, and the disposition to be deliberate about being open-minded or intellectually honest, leading to logical and appropriate action (Huang et al., 2016). Furthermore, Bassham et al. (2018) state that critical thinking is a broad set of cognitive abilities and intellectual characteristics required to properly detect, analyze, and evaluate arguments and truth assertions. Critical thinkers should identify and overcome their prejudices and assumptions, develop and present convincing arguments supporting conclusions, and evolve to logical, informed decisions about what to think and do. Therefore, critical thinking is crucial for effective problem-solving, decision-making, and innovation to helps individuals evaluate information accurately.

The importance of critical thinking makes incorporating it into a curriculum need to be considered. It has been chosen because of its broad relevance to the educational system. The current educational system frequently encourages repeating and memorizing facts, which is knowledge, without actively involving students in the discovery process (Straková & Cimermanová, 2018). It makes students unable to think critically and be actively involved in making decisions. Thus, The Indonesian Ministry of Education organized teacher training to give teachers the necessary understanding of what was being taught following Curriculum 2013, including critical thinking. The programs aimed to increase

teachers' overall understanding of critical thinking. Teachers as facilitators must comprehensively understand the culture of critical thinking to enhance the student's critical thinking ability (Kim & Pollard, 2017). Therefore, critical thinking should be taught to the students.

Students are expected to think critically to solve their problems and decide the appropriate choice. However, students have different ways of producing ideas. A study by Rianto (2021) revealed a significant difference between male and female students in terms of thinking process. Gender affects linguistic indicators in writing, such as males and females having different word choices in expressing their ideas through writing (Pahamzah et al., 2022). Moreover, Hz (2022) argued that males emphasize logic while females are more concerned with social construction. It means that a male is rational, preferring the presented message's importance rather than considering what others may think and feel. Thus, gender is an interesting aspect of critical thinking that must be further examined to know the differences between male and female students in critical thinking and what makes males and females different in generating ideas.

People put their ideas and critical thoughts into writing. In an academic context, the use of argumentative writing is an option. Argumentative writing is a specific form of academic writing that aims to persuade readers to adopt a particular perspective or take a specific course of action on a controversial or debatable topic. It involves presenting a well-structured argument supported by evidence and logical reasoning. It emphasized the importance of clear claim statements, logical organization, and evidence-based reasoning in constructing persuasive arguments (Andrews et al., 2018). Writing an argumentative can promotes critical thinking, analytical skills, and knowledge retention (Graham et al., 2011). It is a powerful tool for expressing opinions, influencing readers, and engaging in intellectual discourse. Constructing a compelling and persuasive argument requires careful analysis, strong reasoning, and effective communication skills. So, argumentative writing can measure students' critical thinking and the differences between male and female students in critical thinking.

Several studies have attempted to investigate the issues around gender and students' critical thinking, especially in writing essays. For example, a study by Preiss, D. D., et al. (2013) entitled *"Assessment of Argumentative Writing and Critical Thinking in Higher Education: Educational Correlates and Gender Differences"* shown the findings of the study demonstrated that gender differences in thinking and writing are essentially independent of similarities. In this study, males outperformed females in ability of thinking which had a good syllogistic reasoning and inference analysis tests. Meanwhile, females outperformed males in ability of writing an argumentative. However, this study had some inconsistencies in the results of argumentative writing measure since the differences between male and female students in critical thinking were different. Then, this study used written communication test which provided an average score of analytical score, but the latter score was holistic score.

The second study by Noroozi et al. (2018) entitled *"Students' Online Argumentative Peer Feedback, Essay Writing, and Content Learning: Does Gender Matter?"* revealed the differences between male and female students in their argumentative feedback on the quality. Male students provided lower-quality argumentative feedback than female students. This research contributed to a growing body of research showing that how students engage in argumentative peer feedback differs between genders, even though gender does not affect how well students write essays or learn about the subject matter. In other instances, we could not identify significant differences between female and male students based on the literature because this study was conducted in a massive power distance community.

The other research conducted by Cáceres et al. (2020) entitled *"Integrating Critical Thinking into The Classroom: A Teacher's Perspective"* highlighted the teacher's perspective as the key to bridging the students' critical thinking skills. Instead of teaching critical thinking abilities separately, teachers aim to help their students acquire critical thinking by including it in their lessons. Students had to analyze, investigate, and develop solutions for most of the issues raised. The findings of this study showed that the teachers concur that critical thinking is more

than a cognitive process or a set of rules; it involves utilizing subject-specific information and practices to address real-world situations. However, this research focused on the teacher's perspective, who integrates critical thinking in discussion or speaking ability.

Marni S. et al. (2020) conducted a study entitled *"Students' Critical Thinking Skills based on Gender and Knowledge Group."* This study showed that 25 statements and four indicators were used to measure students' critical thinking. This study implemented Facione's (2015) theory, classifying the indicators into interpretation, analysis, evaluation, and inference. This study concluded that there were no significant differences between males and females in their critical thinking. However, the total value of each indicator showed male students were higher than female students. As observed from knowledge group in this study, it has shown that there were no significant differences between critical thinking skills of science students and humanity students in general.

Darmaji et al. (2021) conducted a study entitled *Relationship of Science Process Skills on Critical Thinking Ability Review by Gender in Madrasah Aliyah.* This research involved the students at MAN 5 Batanghari. Research shows the science process skill and students' critical thinking skills in terms of gender. It was proved that male and female students are good at critical thinking. Meanwhile, male students were superior in each class to female students in science process skills. This research indicates a strong relationship between science process skills and critical thinking skills. The results of inferential statistical data supported it through the Pearson correlation test, the obtained value of 0.648. However, this research focused on the science process skills that can help to improve students' critical thinking abilities.

This current study is different from five previous studies. This study investigates whether gender has differences in each indicator of critical thinking in argumentative writing. The gender-focused study has not provided strong confirmation of these differences' effects. Thus, this study will examine whether the result will support or disapprove of the gender concerning academics, especially in critical thinking of argumentative writing. Secondly, this research will synthesize

critical thinking aspects (Facione, 2015) and argumentative writing aspects (Heaton, 1988; Toulmin et al., 2002) to measure students' argumentative writing.

1.2 The Research question

1. Is there any significant difference between male and female students in critical thinking in argumentative writing?

1.3 The Objectives of the Study

1. To identify whether there is any significant difference between male and female students in critical thinking in argumentative writing.

1.4 The Significance of the Study

Theoretical aspects. The researcher expects that this research will enrich the existing knowledge about the differences between male and female students in critical thinking through argumentative writing.

Practical aspects. The results of this research will contribute to students at the University of Muhammadiyah Malang in critical thinking skills. This research results can impact teachers' understanding to ensure male and female students are supported in developing their critical thinking skills in argumentative writing. In addition, this study is also expected to benefit English Language Department by facilitating discussions that involves male and female students' viewpoints on their understanding of critical thinking in argumentative writing. Then, the material developers can offer to the students about supplementary resources, alternative examples, or optional activities that comply with different learning styles and preferences.

1.5 The Scope and Limitation

The primary focus of this study is to determine the differences between male and female students in critical thinking through argumentative writing. However, this study involves undergraduate students at the University of Muhammadiyah Malang. Thus, these results may not represent undergraduate students in Indonesia.

1.6 The Definitions of Key Terms.

Gender: Gender is a cultural term connected to biological sex to distinguish the characteristics between males and females (Archer & Lloyd, 2002). In this study,

the researcher will measure significant differences in critical thinking between male and female students based on their argumentative writing scores.

Critical Thinking: Critical thinking is an analysis and consideration of certain situations with value judgments based on the facts-evidence (Moon, 2007). In this study, critical thinking will be judged or determined based on whether or not students' compositions include sufficient data or evidences.

Argumentative Writing: Argumentative writing is a type of discourse that tries to persuade readers by offering well-organized and well-supported arguments based on facts, logic, and critical thinking (Birkenstein & Graff, 2018). In this study, argumentative writing is the research object used to identify whether the essay contains of critical thinking aspects.

CHAPTER II

REVIEW OF RELATED LITERATURE

This chapter discusses three main points. They are relevant theoretical backgrounds related to the current study, conceptual framework and research hypotheses.

2.1. Critical Thinking

2.1.1 Definition of Critical Thinking

Critical thinking is a human way of responding to someone or an article by analyzing facts to assess these facts. Analyzing, evaluating, and considering ideas to improve them is the art of critical thinking; a systematic approach to study and intellectual brilliance are required for critical thinking (Paul & Elder, 2009). It demands that we look beyond our current understanding and consider the advantages of current knowledge and different points of view. Critical thinking is an ability to comprehending, assessing, and evaluating information to judge or determine whether something is right or wrong. In line with Ennis (2011) argued that making decisions about what to believe or do is the main goal of critical thinking, which is a reflective thinking process. Critical thinking is the writer's capacity to analyze their thinking to identify its advantages and disadvantages. To present correct judgment, reasoning, and comprehension during critical thinking, we require access to enough information about a topic or idea. Consequently, change your perspective and restructure your thoughts to be more effective; the intellectual skill of critical thinking will help you in both your professional and educational pursuits. It must support their stance within the critical thinking process by offering evidence regarding the topic they are addressing so that their conclusions are perceived as solid and proven (Judge et al., 2009).

Critical thinking skill is broad and encompasses a variety of desirable results, such as good professional knowledge, which brings us to the conclusion that it is linked to various other desirable professional goals and traits. It's a current issue since failing to draw the appropriate distinctions will prevent critical thinking

from developing properly (Penkauskienė et al., 2019). Therefore, critical thinking is commonly acknowledged as one of the twenty-first century's essential abilities. Asia has several nations that have incorporated critical thinking skills into their educational reforms, including Singapore and Hong Kong (Mok, 2011), China (Lin, 2018), and Iran (Afshar & Movassagh, 2017). In numerous policies in Indonesia, critical thinking is required (see MoEC Decree No.17/2010, Article 77, and MoEC Decree No.23/2016). Critical thinking is the ability to objectively and systematically analyze and evaluate information, arguments, and situations logically and systematically. It involves questioning assumptions, examining evidence, considering different perspectives, and drawing reasoned conclusions. It should be directed toward improved inquiry into questions of truth and untruth and issues of meaning more generally, as well as imperatives and possibilities of moral and political action.

Critical thinking enables the students to evaluate the facts in light of what they have read and to spot incorrect or illogical thinking (Keynes, 2008). Additionally, it will help students when using critical thinking to strengthen arguments (for instance, in assignments). It entails seeing and defending any claims made in the context of the examined evidence. The evidence discussed in the work is founded on facts rather than the author's opinions. Therefore, critical thinking is intimately tied to argumentative essays because it requires the writer to analyze, compile, and present the arguments in each current work (Tilaar, 2011). Thus, Johnson E. (2006) argued that students who possess strong critical thinking abilities are bound to be able to analyze a topic in the present with accuracy, deal with the millions of problems that arise systematically, ask novel questions, and come up with solutions that are still relatively novel.

2.1.2 The Characteristics of a Critical Thinker

Critical thinkers have several characteristics that allow them to approach making opinions with evidence, problem-solving skills, and decision-making with several logical assumptions and premises. It was in line with Bassham et al. (2011), that offered some general profiles of critical thinkers:

1) A critical thinker has an intense need for correctness, precision, and other aspects of critical thinking.
2) A critical thinker is competent at comprehending, analyzing, and classifying claims and points of view.
3) A critical thinkers use logic to justify reasoning and use evidence and data to draw proper conclusions.
4) A critical thinker is willing to consider opposing viewpoints with an open mind and appreciates criticisms of their views and assumptions.
5) A critical thinker can cut through the clutter of facts and reach the core of a situation or issue.
6) A critical thinker has the strength to consider truthfully evaluate arguments that contradict even their most fundamental assumptions.
7) A critical thinker has to build opinions on information and facts rather than bias or self-interest.
8) A critical thinker is conscious of the assumptions and biases that influence their perception of facts.
9) A critical thinker is a person of integrity who isn't afraid to question what is commonly believed.
10) A critical thinker is not afraid to disagree with a group or others as long as they have strong evidence, facts, and assumptions.

The characteristics shown above are that a good critical thinker must be able to give assumptions, solutions, and criticize some issues with strong facts and evidence. Thus, a critical thinker needs to fulfil their critical thinking ability by knowing the aspects of critical thinking.

2.1.3 Aspects of Critical Thinking

Students or someone else has to be able to think critically, and they should learn how since these abilities are essential for navigating life both now and in the future. Thus, students with critical thinking capacity can process information

logically and rationally and solve difficulties methodically; it is essential to understand some aspects that make it up. All aspects of critical thinking are interconnected and do not differ; each argument must be supported by relevant facts and data (Zakiah & Lestari, 2019). Then, Garrison et al. (2011) argued that aspects of critical thinking are divided into four aspects: identifying or acknowledging a problem; deciding in advance after giving personal and societal concepts some thought; integration, or creating the idea's meaning; offering speculative solutions or implement actual solutions to problems. It was in line with Facione et al. (2015) stated that critical thinking is the act of using to make a deliberate, and purposeful conclusion and the experts identified a set of aspects of critical thinking which classified into interpretation, analysis, inference, explanation, evaluation, and self-regulation.

1) Interpretation refers to the act of conveying the significance or relevance of various experiences, circumstances, data, events, judgments, norms, beliefs, rules, procedures, and criteria.
2) Analysis is the act of identifying the intended and actual logical connections between statements, questions, concepts, descriptions, or other forms of representation that convey belief, judgment, experiences, reasoning, facts, or opinions.
3) Inference is the cognitive process of figuring out logical conclusions by analysing and interpreting relevant information, such as data, statements, evidence, or beliefs. It involves forming conjectures and hypotheses, as well as considering the implications and consequences of the information at hand.
4) Explanation entails the process of providing a rationale for one's reasoning by considering facts, conceptual aspects, methodology, criteria, and contextual factors, resulting in persuasive arguments.
5) Evaluation is the procedure of assessing the reliability of statements or other descriptions that depict an individual's viewpoint, encounter,

circumstance, judgment, conviction, or opinion. It assesses the logical soundness of the existing or intended connections between claims, descriptions, queries, or other types of representation.

6) Self-regulation entails the process of critically examining and assessing one's own inferential judgments in order to evaluate, affirm, authenticate, or rectify thinking and outcomes.

As mentioned, these six aspects may be a foundation for developing critical thinking, eventually resulting in a strong argument. A strong argument represents something that is supported by true claims and facts.

2.1.4 Benefits of Critical Thinking

After discussing some aspects of intellectual rules that guide critical reasoning, we will investigate the benefits of critical thinking. Bassham et al. (2018) classified the benefits of critical thinking into three levels:

1) Benefits of Critical Thinking in the Classroom: It focused on college students. Critical thinking is a crucial function throughout the college curriculum. Thus, students can identify other people's ideas and viewpoints. Moreover, they can consider those claims and arguments to create well-supported ideas. It's helpful for college students since critical thinking skills are typically required for college tasks alongside scientific papers.

2) Benefits of Critical Thinking in the Workplace: workers with good critical thinking problem-solve, analyze some information, and draw conclusions from the data. These are the kind of critical thinking skill that employees value and is crucial in the workplace.

3) Benefits of Critical thinking in life: critical thinking is worthwhile to study simply for the personal enrichment and empowerment it can provide to our life; we can prevent making poor decisions for ourselves by using critical thinking. Critical thinking is essential for advancing democratic processes.

Moreover, Critical Awareness enhances our critical ability, allowing us to discover areas that indicate applying more systematic criticism during our thinking and behaviours. Cottrell (2023) classified the benefits of having good critical thinking into twelve benefits:

1) Capability to identify the assumptions.
2) Capability to identify issues and potential problems that require additional investigation.
3) Capability to make wise decisions.
4) There is less chance of being deceived.
5) Capability to recognize what is essential to save time.
6) Capability to improve accuracy and precision in several aspects of a task.
7) Better communication and thinking.
8) Improved problem-solving abilities, such as identifying areas for improvement and analyzing possible solutions.
9) The ability to use a systematic approach ensures that basics are not ignored.
10) Increased speed and accuracy in analyzing difficult data.
11) Self-assurance in confronting more complex problems and tasks.
12) Possibility of perceiving the world with eyes and greater awareness.

2.1.5 Developing Critical Thinking

Gaining critical thinking skills has several benefits. As mentioned before, critical thinking benefits consist of improved problem-solving abilities, a lower chance of being duped or deceived, increased confidence, and increased speed and accuracy while analyzing complicated information. Therefore, the students must be aware and expected to develop their critical thinking ability (Cottrell, 2023). These are the ways to develop critical thinking:

1) Find out where the best evidence lies for the subject you are discussing.
2) Evaluating the strength of the evidence to support different arguments.
3) Coming to an interim conclusion about where the available evidence appears to lead.
4) Constructing a line of reasoning to guide your audience through the evidence and lead them towards your conclusion.
5) Selecting the best examples.
6) Provide evidence to illustrate your argument.

2.2. Gender and Thinking Style

Gender describes the socially built characteristics distinguishing males, females, girls, and boys. It was in line with Mosse (2003), who argued that gender is a grouping of roles that allow people to identify as either male or female to others. As a social construct, gender is viewed differently in many countries. Significant gender disparities exist in every society; they significantly affect people (Talbot, 2010). Because of the social and cultural differences in the features that exist in males and females, gender can be studied (Mosse, 2003; Talbot, 2010). There are three distinct gender understandings, such as gender roles, gender typing, and gender identity (Santrock, 2013). Gender roles are expectations for how males and females should feel, act, and think. Most kids act in ways that increasingly reflect the gender norms prevalent in their culture. Gender typing refers to taking on a conventionally masculine or feminine role. It was in line with Golombok et al. (2008), who argued Fighting is a typically male role reference, whereas crying is a feminine role; this sort of conduct indicates sex-typed behaviour. Gender identity refers to a person's view of one's gender, covering knowledge, comprehension, and acceptance of being both male and female (Egan & Perry, 2001; Perry, 2012). Then, Blakemore et al. (2009) argued one aspect of gender identification is knowing whether you are male or female, and most children understand this concept by the time they are $2^{1/2}$ years old.

Gender-specific learning and living styles result from differences in the structure and function of male and female brains, not environmental factors. The brains of most females mature more quickly and early than those of most males. Girls may develop complex linguistic skills up to a year earlier than boys (Gurian & Ballew, 2003). According to gender-with-achievement relationships, males achieve slightly more in science than girls. Men scored higher than women according to gender and cognitive ability (Steinkamp & Maehr, 1983). It was in line with Lever (1978), who demonstrated that the structure of males' play is more advanced than that of girls' play. Boys' play is often characterized by specialized roles, player interdependence, clear group goals, more significant group sizes, many regulations, and team divides. Girls' play is generally less motoric and manipulative than boys' (Lewis, 1972). Gender discrimination exists in the field of education and has an impact on the learning process (Darmaji et al., 2021). It's an important issue that has been hugely overlooked about gender and education. It's recognized that gender matters in education. In recent years, there has been a lot of study on how gender influences first language (L1) writing performance (Al-Saadi, 2020).

Islam (2018) claims that incorporating gender topics into ELT classes, particularly those in higher education, may raise awareness of women's rights and empowerment; he continues by saying that gender equality and a hostile environment are essential for any democratic society to succeed. There are few studies on EFL students' difficulties writing essays on gender topics. The following factors make it vital to raise gender concerns when writing an essay. Firstly, Indonesia continues to have gender disparity. Gender disparities in a variety of aspects of life are evidence that gender equality and women's empowerment in Indonesia are still in need of improvement, according to data from the Directorate of Demography, Women Empowerment, and Child Protection of Indonesia (The Ministry of Demography, Women Empowerment, and Child protection of Indonesia/Bappenas, 2014). Secondly, both the lecturer and the students can benefit from a knowledge of students' difficulties while writing essays, particularly those that encourage higher-order thinking, like an argumentative essay (Setyowati et al., 2020). Alongside argumentative essays, gender is a crucial argumentative

consideration (Asterhan et al., 2012; Tsemach & Zohar, 2021). Furthermore, there are differences in how male and female students handle peer review and essay writing in online learning contexts. It confirmed the evidence that suggests that gender significantly influences how well argumentative essays are written and peer reviews are performed (Noroozi et al., 2022).

2.3. Argumentative Writing

2.3.1 The Concept of Argumentative Writing

Writing is a semiotic tool that promotes social interaction and communication, is learned and utilized in social contexts, and is employed to achieve inherently social aims (Bazerman, 2016; Graham, 2018; Newell et al., 2018). Argumentative writing defends an opinion but does not attempt to persuade the reader's emotions. It occurs when you have a message, i.e., something to say or a point to make, and due to your original thought, it is then possible for there to be interaction between various viewpoints on this message. The current theories of argumentative writing (Ferretti & Fan, 2016) acknowledge its innately social and dialogical nature and that it entails conveying a set of arguments meant to accomplish the deliberate goals of the counterparts (van Eemeren, 2018). Argumentative writing is regarded as a crucial educational goal and a well-liked activity for college students (Asterhan, 2018; Noroozi et al., 2016; Wu, 2006), particularly when they tackle challenging and divisive topics.

Meanwhile, Langan (2007) stated that the main purpose of argumentative writing is to persuade the audience that the writer's particular viewpoint or opinion on a certain topic is true and to move the audience to take some action. A writer of argumentative writing tries to defend opposing perspectives with reasoned arguments and ideas grounded in reality (Lai, 2011). Thus, argumentative writing requires the student to analyze an issue, gather, produce, and assess supporting data, and clearly state their point of view. It was in line with Wentzel's (2018) statement that argumentative writing involves more than merely expressing your views or criticizing others. Furthermore, Ferretti & Graham (2019) highlighted the issue of

various theoretical viewpoints and analytical techniques applied to study argumentative writing and comprehend the circumstances that affect its growth. Our comprehension of argumentative writing has significantly benefited from social, cognitive, and linguistic viewpoints.

Thus, students often hesitate to utilize argumentation strategies in their written argumentative essays and various argumentation problems in various courses, and students frequently struggle when preparing these essays (Noroozi et al., 2018; Wingate, 2012). Most of the students struggled to write argumentative writing because they did not know the aspects of argumentative writing. It was in line with Kuhn (1991), who argued that argumentative writing is a critical thinking technique for developing ideas, solving problems, and exercising sound judgment. Essentially, the students need to master in writing argumentative writing, which should have good arguments in conveying it; good arguments have two sides in developing an argumentative writing essay that has two sides: claims and counterclaims (Nussbaum & Schraw, 2007). Integration of claims and counterclaims is regarded to be more believable in written texts since the author comes off as more educated and less prejudiced (O'Keefe, 1999).

2.3.2 Aspects of Argumentative Writing

Argumentative writing has several aspects, each serving a specific role in building a good argument. Claims and counterclaims are the aspects of argumentative writing (Nussbaum & Schraw, 2007). Meanwhile, according to Toulmin et al. (2002), the primary aspects of argumentative writing include:

1) Claims refer to an opinion or argument as an assertion to certain issues, and the first step in analyzing and criticizing the argument is to understand the precise issues.

2) Grounds, which form the foundation of an idea known as the premise, can indicate the acceptance and accuracy of a claim. It may establish its truth. Correctness or soundness includes observations, common knowledge, statistical evidence, personal testimony, already established

assertions, or other "factual data" that can support claims.

3) Warrants mean guarantee. It refers to the facts or supporting evidence from the argument. Ground is a foundation of claims that relies on the facts, and warrants support the ground and claims by giving more evidence or facts that have been made.

4) Backings refer to any additional support of the warrant. It supports the warrant by giving a specific example that justifies it.

5) Modal Qualifier allows for limiting the limitation of the claim and indicates an amount of confidence in the conclusion made from the facts in the warrant.

6) The rebuttal is the recognition of opposing perspectives. It will be beneficial to persuade hesitant audience members to concur with the statement or claim.

Therefore, claims made in arguments are solidly supported if the grounds are correctly provided. These grounds must be connected to the claims by valid, applicable warrants that can, in turn, be supported by enough backings of the appropriate kind.

2.4 Conceptual Framework

Gender is classified into male and female, which can produce critical thinking. Males and females have different ways of conveying critical thinking into some writing products. The writing product that has critical thinking aspects is argumentative writing. Then, male and female students must have the ability to use critical thinking to provide good argumentative writing. As a result, a researcher will reveal the differences between male and female students in critical thinking in their argumentative writing. They are supposed to have good critical thinking skills to write an argumentative essay. Meanwhile, if they have poor critical thinking, they are supposed to have poor argumentative writing.

2.5 Research Hypotheses

This study proposes some hypotheses as follows:

1. Null hypothesis (H0): there is no significant difference between male and female students in critical thinking in argumentative writing at the University of Muhammadiyah Malang.
2. Alternative hypothesis (Ha): there is significant difference between male and female students in critical thinking in argumentative writing at the University of Muhammadiyah Malang.

CHAPTER III

RESEARCH METHOD

In this chapter, the researcher discusses five main points used to conduct this study. They are research design, research subject, population & sample, data collection, and data analysis.

3.1. Research Design

This research used an ex post facto design with a quantitative descriptive method. Similarly, researchers occasionally use "ex post facto" to describe causal-comparative research (Fraenkel et al., 2022; Ravid, 2020). When a researcher cannot control the treatment variable, they use the ex post facto design, which involves collecting a post-test without any pre-tests to describe current events (Mertler, 2020; Ravid, 2020; Edmonds & Kennedy, 2017; Kathori, 2004). Thus, in this study, the researcher used the existing students' argumentative essays to analyze and score by two raters. The researcher collected the existing students' argumentative essays from the lecturer who taught the essay writing course with the aim of empowering students to concentrate on their writing and minimize their inattention, ensuring the achievement of all existing indicators. The researcher synthesized the indicators of critical thinking (Facione, 2015) and argumentative writing (Heaton, 1988; Toulmin et al., 2002) to score the students' essays. Then, the indicators of critical thinking (Facione, 2015) consist of six indicators: interpretation, analysis, inference, explanation, evaluation, and self-regulation. They have theoretical similarities with the indicators of argumentative writing (Toulmin et al., 2002). It comprises six indicators: claims, grounds, warrants, backings, modal qualifiers, and rebuttals.

Furthermore, the primary benefit of descriptive statistics (Fraenkel et al., 2022) is that quantitative descriptive analysis enabled the researchers to effectively describe the information from numerous scores using only a few indices, such as the mean and median (more on these in a moment). Then, this research used the notions of explanatory and prediction research design to determine the significant

differences between male and female students, explain, and interpret the results. Thus, the data of the students' scores was analyzed by using JASP and SPSS to determine the differences between male and female students in critical thinking. Moreover, the variables are classified into independent and dependent variables (Fraenkel et al., 2012). This study examined the independent variable (X) that refers to gender (male and female students at the University of Muhammadiyah Malang). Meanwhile, the dependent variable in this study is students' critical thinking (Y), which refers to the indicator used to measure students' critical thinking in argumentative writing. This research involved finding out the significant differences between male and female students in critical thinking. Furthermore, the researcher interpreted the results. The variables of this research design were as follows:

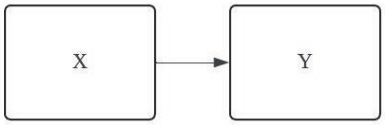


Figure 1. Variables

X = Gender (Male and Female)

Y = Students' critical thinking

3.2 Data Collection

To collect the data, the researcher collected the students' argumentative writing essays from the lecturer who taught the essay writing course. The researcher in this study collected 42 students' essays, which consisted of 26 female students and 16 male students. In this study, proportional stratified sampling was used to select student essays. However, obtaining a sufficient number of student essays is not feasible due to their small size. When proportional stratified sampling is not possible for any reason, the study can use quota sampling (Salkind, 2012). In addition, stratified random sampling is similar to quota sampling, which is a non-probability method that divides a sample into groups based on characteristics like age and gender (Lohr, 2010; Gorny & Napierala, 2015). This aligns with Fraenkel et al.'s (2022) statement that causal-comparative or ex post facto research requires a minimum of 15 participants per group, or at least two groups. Therefore, the

researcher took into account the total number of student essays, specifically including 15 essays from female students and 15 from male students. Thus, this study used quota sampling, which consisted of two groups, such as male and female students. The study specifically focuses on the following characteristics of the sample:

1. The English Language Education Department at the University of Muhammadiyah Malang.
2. Students' argumentative essays.
3. 15 male and 15 female students' essays.

Furthermore, the researcher sent the synthesized indicators to the raters to score the essays, as well as the students' argumentative essays. The rater's scoring of an individual's conduct or product informs the design of the ratings. Fraenkel et al. (2022) also classified the rating scales into two categories: behavior rating scales and product rating scales; they identified common educational ratings to include book reports, essay writings, diagrams, drawings, and notebooks. Thus, this study used a rating with an ordinal scale to score the students' essays.

3.3 Data Analysis

Descriptive statistics are statistical methods used in data analysis that provide a description or visual representation of the acquired data. Similarly, Fraenkel (2012) argued that descriptive statistics enable researchers to summarize a large amount of data using a few key measures, such as the mean, median, and standard deviation. Thus, this research used descriptive statistics. Then, the data analysis in this research involved three steps: the validity test, the reliability test, and the Mann-Whitney test. Validity is centered on the defensibility of the findings that researchers draw from the data collected through the use of instruments. Therefore, a valid instrument is one that measures the intended object (Fraenkel et al., 2022). Thus, this research involved two validators to validate the instrument of a critical thinking indicator for scoring students' argumentative writing. Furthermore, this study involved two raters to score the students argumentative

writing, which determined reliable data by the inter-rater reliability test. An instrument that consistently produces results is considered reliable (Fraenkel et al., 2022).

Furthermore, the researcher utilized Microsoft Excel 2017 to categorize the scores of male and female students, each critical thinking indicator. Then, inter-rater reliability provided an overview in the form of a score regarding the extent of agreement given by the raters or experts using Cohen's Weighted Kappa. The researcher calculated the inter-rater reliability using the JASP interface (Jeffreys's Amazing Statistics Program). The JASP interface was more like SPSS, but JASP enabled the researcher to conduct statistical analyses with simplicity in mind. Furthermore, the researcher used SPPS for the normality test to determine whether the data is normally or non-normally distributed, as well as the Mann-Whitney U test to determine the significance difference between male and female students in critical thinking.

3.3.1 Validity Test of Indicator Critical Thinking for Argumentative Essays

This study validated the critical thinking indicators to score students' argumentative writing using face validity and content validity. Face validity refers to the extent to which examinees believe the instrument is measuring what it is supposed to measure (Ary et al., 2019). This research used face validity and content validity as indicators of critical thinking to score students' argumentative writing. Meanwhile, content validity refers to the degree to which a research instrument accurately measures all indicators. The researcher used critical thinking (Facione, 2015) and argumentative writing (Heaton, 1988; Toulmin et al., 2002) as indicators for scoring argumentative essays in this research. This approach provides students with writing opportunities that mirror the real-world skills required by reflecting the authenticity of the target contexts (Hyland, 2003). Then, the researcher sent the indicator for scoring argumentative writing to the validators or experts by email. Next, the researcher asked the experts to analyze the proposed aspects of critical thinking and argumentative writing. Their judgments show that the synthesized indicators of critical thinking (Facione, 2015) and argumentative writing (Heaton,

1988; Toulmin et al., 2002) were appropriate. In choosing the validators or experts, the following criteria were considered:

1. The validators had ≤ 5 years of teaching experience.
2. The validators had teaching experience in English or Writing courses.
3. The validators were Doctoral graduates from the English Language and Education Study Program.

3.3.2 Reliability Test of Argumentative Essay Score

This study used inter-rater reliability, which allowed the raters to give a score or rating to the students' essays, whether it was appropriate for the scoring indicator or not. Similar to Douglas (2010), who stated that to evaluate speaking and writing ability, reliability can be divided into two main categories: intra-rater reliability and inter-rater reliability, Intra-rater reliability is defined as how consistently a rater assigns the same score to a given performance. Meanwhile, inter-rater reliability refers to how consistently two raters assign the same score. The researcher sent the documents for the indicator for scoring argumentative writing and the students' essays to the inter-raters by email. Then, the two raters were asked to score the students' essays using the indicators given. Their judgments showed that the students' scores were according to the synthesized indicators of critical thinking (Facione, 2015) and argumentative writing (Heaton, 1988; Toulmin et al., 2002). In choosing the inter-raters or experts, the following criteria were considered:

1. The raters had ≤ 3 years of teaching experience.
2. The raters had teaching experience in English or Writing courses.
3. The raters were Magister graduates from the English Language and Education Study Program.

Furthermore, this study utilized Coefficient Cohen's Kappa for inter-rater reliability which Cohen's kappa is specifically intended to measure agreement between two raters (Cohen, 1968). Table 3.2 presents the Cohen's Kappa values, potential degrees of inter-rater reliability, and the percentage of dependability utilized in this research for inter-rater reliability.

Table 3.1 Interpretation Score of Coefficient Cohen's Kappa

Score of Coefficient Cohen's Kappa	Level of Reliability	Percentage of Reliability
< 0.00	Poor Agreement	0 – 4%
0.00 - 0.20	Slight Agreement	4 – 15 %
0.21 - 0.40	Fair Agreement	15 - 35%
0.41 - 0.60	Moderate Agreement	35 – 63%
0.61 - 0.80	Substantial Agreement	64 – 81%
0.81 - above 0.90	Almost Perfect Agreement	82 – 100%

Source: (Landis, & Koch., 1977; McHugh, 2012)

3.3.3 Normality Test

A fundamental assumption in classical statistical analysis, the normality test assesses whether a dataset follows a normal distribution. This test also plays a crucial role in determining the quality of the data, enabling further data analysis procedures to proceed. Therefore, this study used data of no more than 50, while the total in this research is 30. To determine the distribution of random data from a small sample, the researcher used the Shapiro-Wilk normality test. Interpret the Shapiro-Wilk normality test findings by examining the significance value. On the other hand, if the p-value, or Sig., is larger than 0.05, the data is determined to be normally distributed. Shapiro & Wilk (1965) determined that the data is non-normally distributed if the p-value, or (Asymp.) Sig. is less than 0.05. It can be described as follows:

a) If the p-value (Asymp. Sig.) < 0.05, then data is non-normally distributed.

b) If the p-value or (Asymp. Sig.) > 0.05, then data is normally distributed.

3.3.4 The Mann-Whitney U Test

Conceptually, the Mann-Whitney U test and the t-test are similar in that they both determine whether two sampled groups belong to the same population. The Mann-Whitney test is a measurement scale with ordinal data. Moreover, the Mann-Whitney U test employs statistical analysis to distinguish between two groups using a single ordinal variable, without assuming any assumptions about the data's

distribution (Mann & Whitney, 1947; Conover, 1999; Sheskin, 2011). Then, Conover (1999) also stated that the Mann-Whitney U test determined the existence of a significant difference between two independent samples. Thus, this study has ordinal data, and the sample size is small. The researcher uses a Mann-Whitney test to differentiate the critical thinking score between male and female students through their argumentative essays. The following examples describes the decision criteria of the Mann-Whitney test:

a) If the Asymp. Sig. (2-tailed) < 0.05, then there is a significant difference between male and female students in critical thinking.
b) If the Asymp. Sig. (2-tailed) > 0.05, then there is no significant difference between male and female students in critical thinking.

CHAPTER IV

FINDINGS AND DISCUSSION

This chapter will present the findings and discussion. The findings comprise inter-rater reliability, a statistical description, the Mann-Whitney test, and a discussion that includes an analysis of the findings.

4.1 Findings

In this study, the researcher uses quantitative data analysis to investigate the differences between male and female students in critical thinking approaches to students' argumentative essays. The researcher analyzes the students' essay scores using three test statistics, as follows:

1) Inter-rater reliability test is used to measure the agreement between two raters. The results of the inter-rater reliability test were interpreted by the Coefficient Cohen's Weighted Kappa *(Chapter 3/Table 3.1/p. 24).*
2) The normality test is used to determine the quality of data that is normally distributed or non-normally distributed. This study used data from 30 students' essays. Then, the researcher read the statistical data from the Shapiro-Wilk normality test and interpreted it as follows: if the p-value (Asymp. Sig.) is less than 0.05, it indicates a non-normal distribution, and if it is higher than 0.05, it indicates a normal distribution.
3) The Mann-Whitney U Test is used to differentiate between male and female students in critical thinking through their argumentative essays. The Mann-Whitney test divides its structure into six indicators of critical thinking: interpretation (claims), analysis (grounds), inference (warrants), explanation (backings), evaluation (modal qualifiers), self-regulation (rebuttals), and the critical thinking score. The researcher then explains the decision criteria of the Mann-Whitney test to determine a definitive conclusion about the difference between male and female students in critical thinking. If the Asymp. Sig. (2-tailed) is less than 0.05, there is a significant difference between male and female students in

critical thinking. Otherwise, if the Asymp. Sig. (2-tailed) is greater than 0.05, there is no significant difference between male and female students in critical thinking.

Furthermore, the researcher conducts the three statistical tests mentioned above gradually and sequentially, as explained below.

Inter-rater Reliability Test

The researcher uses an inter-rater reliability test to measure the agreement between two raters. In this study, the researcher reveals the agreement between two raters in six critical thinking indicators. The six critical thinking indicators consist of interpretation (claims), analysis (grounds), inference (warrants), explanation (backings), evaluation (modal qualifiers), and self-regulation (rebuttals). This process continues with the overall score for critical thinking. Tables 4.1 explain the inter-rater reliability results from the six indicators of critical thinking.

Table 4.1: Inter-rater Reliability Results from Six Indicators of Critical Thinking

Six Indicators	Cohen's Weighted kappa First Rater – Second Rater
Interpretation (Claims)	0.874
Analysis (Grounds)	0.897
Inference (Warrants)	0.816
Explanation (Backings)	0.893
Evaluation (Modal Qualifiers)	0.741
Self-regulation (Rebuttal)	0.769

The inter-rater reliability results in Table 4.1 show Cohen's weighted kappa of the six indicators of critical thinking. The first interpretation (claims) indicator has a weighted kappa value of 0.874 *(Chapter 4/Table 4.1/p. 27),* and the interpretation score of Cohen's Weighted kappa *(Chapter 3/Table 3.1/p. 24)* shows

an almost perfect level of agreement, with a reliability range of 82% to 100%. *As a result, the researcher declared the interpretation (claims) indicator to be reliable.*

The second indicator is analysis (grounds) indicator. The analysis (grounds) indicator has a weighted kappa value of 0.897 *(Chapter 4/Table 4.1/p. 27)*, and the Interpretation Score of Cohen's Weighted kappa *(Chapter 3/Table 3.1/p. 24)* shows an almost perfect level of agreement, with a reliability range of 82% to 100%. As a result, *the researcher declared the analysis (grounds) indicator to be reliable.*

Then, the third is the inference (warrants) indicator. The inference (warrants) indicator has a weighted kappa value of 0.816 *(Chapter 4/Table 4.1/p. 27)*, and the Interpretation Score of Cohen's Weighted kappa *(Chapter 3/Table 3.1/p. 24)* shows an almost perfect level of agreement, with a reliability range of 82% to 100%. As a result, *the researcher declared the inference (warrantss) indicator to be reliable.*

Then, the explanation (backings) indicator has a weighted kappa value of 0.893 *(Chapter 4/Table 4.1/p. 27)*, and the Interpretation Score of Cohen's Weighted Kappa *(Chapter 3/Table 3.1/p. 24)* shows an almost perfect level of agreement, with a reliability range of 82% to 100%. As a result, *the researcher declared the explanation (backings) indicator to be reliable.*

Furthermore, the evaluation (modal qualifiers) indicator has a weighted kappa value of 0.741 *(Chapter 4/Table 4.1/p. 27)*, and the Interpretation Score of Cohen's Weighted Kappa *(Chapter 3/Table 3.1/p. 24)* shows a substantial level of agreement, with a reliability range of 64% to 81%. As a result, *the researcher declared the evaluation (modal qualifiers) indicator to be reliable.*

In addition, the self-regulation (rebuttals) indicator has a weighted kappa value of 0.769 *(Chapter 4/Table 4.1/p. 27)*, and the Interpretation Score of Cohen's Weighted Kappa *(Chapter 3/Table 3.1/p. 24)* shows a substantial level of agreement, with a reliability range of 64% to 81%. As a result, *the researcher declared the self-regulation (rebuttals) indicator to be reliable.*

After conducting an inter-rater reliability test, the researcher concluded that the six indicators are reliable. Furthermore, the percentages of reliability for the six indicators and critical thinking are as follows: Four indicators {interpretation (claims); analysis (grounds); inference (warrants); explanation (backings)} are around 82–100%, two indicators {evaluation (modal qualifiers); self-regulation (rebuttals)} are around 64–81%.This percentage does not fall below the lowest inter-rater reliability level of agreement *(Chapter 3/Table 3.1/p. 24)*, which indicates a high level of agreement between the two raters. The researcher concluded that the two raters have reached an agreement to score the students' essays by the instrument of critical thinking scoring, and one of the rater's instrument scores can be representative of the normality test and the Mann-Whitney test for this study. Then, this study uses the first rater's score as an instrument for both the normality test and the Mann-Whitney test, indicating a high level of agreement. The researcher chooses the first rater because he has a longer teaching period and is more qualified than the second rater in terms of writing. However, both of the two raters are lecturers in English language education. *As a result, the six indicators and critical thinking indicator are reliable. Then, the researcher uses the first rater's instrument score, and the instrument can be used to examine the normality test and Mann-Whitney test on the six indicators of critical thinking as well as the students' critical thinking scores.*

Normality Test

The researcher uses the normality test to determine the quality of data that is normally distributed or non-normally distributed. Then, the researcher uses the Shapiro-Wilk normality test because the data involved 30 students' essays. The researcher selects the first rater's instrument for the normality test, focusing on the six indicators: interpretation (claims); analysis (grounds); inference (warrants); explanation (backings); evaluation (modal qualifiers); self-regulation (rebuttals); and critical thinking score to determine the distribution of the score data. However, this is what the inter-rater reliability results showed. The Shapiro-Wilk normality test will interpret it as follows: if the p-value, or Sig., is less than 0.05, it indicates

a non-normal distribution, and if it is higher than 0.05, it indicates a normal distribution. Furthermore, the normality test results presented in *(Appx. 4).*

The normality test result for the interpretation (claims) indicator shows *(Appx. 4)* that the Sig. or p-value is 0.002 for male students and 0.001 for female students, which indicates the Sig. or p-value less than 0.05. *As a result, the normality test for interpretation (claims) is a non-normal distribution.*

Then, the normality test result for the analysis (grounds) indicator shows *(Appx. 4)* that the Sig. or p-value is 0.017 for male students and 0.002 for female students, which indicates the Sig. or p-value less than 0.05. *As a result, the normality test for analysis (grounds) is a non-normal distribution.*

The normality test result for the inference (warrants) indicator shows *(Appx. 4)* that the Sig. or p-value is 0.000 for male students and 0.056 for female students, which indicates the Sig. or p-value of male students less than 0.05, while the Sig. or p-value of female students is higher than 0.05. However, if one group is normally distributed and another is not, the data is not normally distributed. *As a result, the normality test for inference (warrants) is a non-normal distribution.*

Furthermore, the normality test result for the explanation (backings) indicator shows *(Appx. 4)* that the Sig. or p-value is 0.037 for male students and 0.042 for female students, which indicates the Sig. or p-value less than 0.05. *As a result, the normality test for explanation (backings) is a non-normal distribution.*

The normality test result for the evaluation (modal qualifiers) indicator shows *(Appx. 4)* that the Sig. or p-value is 0.001 for male students and 0.056 for female students, which indicates the Sig. or p-value of male students less than 0.05, while the Sig. or p-value of female students is higher than 0.05. However, if one group is normally distributed and another is not, the data is not normally distributed. *As a result, the normality test for evaluation (modal qualifiers) is a non-normal distribution.*

The normality test result for the self-regulation (rebuttals) indicator shows *(Appx. 4)* that the Sig. or p-value is 0.016 for male students and 0.004 for female students, which indicates the Sig. or p-value is less than 0.05. *As a result, the normality test for self-regulation (rebuttals) is a non-normal distribution.*

In addition, the normality test result for the critical thinking shows *(Appx. 4)* that the Sig. or p-value is 0.195 for male students and 0.027 for female students, which indicates the Sig. or p-value of male students higher than 0.05, while the Sig. or p-value of female students is less than 0.05. However, if one group is normally distributed and another is not, the data is not normally distributed. *As a result, the normality test for critical thinking is a non-normal distribution.*

Following the normality test, the researcher comes to the conclusion that the six signs of critical thinking—interpretation (claims), analysis (grounds), inference (warrants), explanation (backings), evaluation (modal qualifiers), self-regulation (rebuttals), and critical thinking—are not normally distributed. Then, the research continues to analyze the instrument's score using the Mann-Whitney test to know whether there is a significant difference between male and female students in critical thinking.

The Mann-Whitney U Test

The researcher in this study uses the Mann-Whitney U test to determine the difference between male and female critical thinking. Then, the researcher also uses the Mann-Whitney U Test to differentiate between male and female in six indicators of critical thinking, such as interpretation (claims), analysis (grounds), inference (warrants), explanation (backings), evaluation (modal qualifiers), and self-regulation (rebuttals). However, before the researcher uses the Mann-Whitney U test, the researcher reviews the descriptive statistics in the table, which consist of the mean, median, and standard deviation. Although the mean difference value is observable and readable, it does not allow for immediate conclusions. As a result, the Mann-Whitney test can strengthen and highlight the differences between males

and females in a statistically significant way. Also, these results are organized around the six indicators of critical thinking which are interpretation (claims), analysis (grounds), inference (warrants), explanation (backings), evaluation (modal qualifiers), and self-regulation (rebuttals). Critical thinking is at the end of list. *(Appx. 5)* and Table 4.2 present the results of descriptive statistic.

The descriptive statistic in *(Appx. 5)* explains the differences in average scores between male and female students for the interpretation (claims) indicator. Male students have an average value of 2.67, and female students have an average value of 2.27. This demonstrates that male students have the highest average score. *As a result, male students outperformed female students on the interpretation (claims) indicator.*

Then, the descriptive statistic in *(Appx. 5)* shows the differences in average scores between male and female students for the analysis (grounds) indicator. Male students have an average value of 2.33, and female students have an average value of 2.27. This demonstrates that male students have the highest average score. *As a result, male students outperformed female students on the analysis (grounds) indicator.*

The descriptive statistics in *(Appx. 5)* shows the differences in average scores between male and female students for the inference (warrants) indicator. Male students have an average value of 2.40, and female students have an average value of 2.27. This demonstrates that male students have the highest average score. *As a result, male students outperformed female students on the inference (warrants) indicator.*

Furthermore, the descriptive statistics in *(Appx. 5)* shows the differences in average scores between male and female students for the explanation (backings) indicator. Male students have an average value of 2.27, and female students have an average value of 2.60. This demonstrates that female students have the highest

average score. *As a result, female students outperformed male students on the explanation (backings) indicator.*

The descriptive statistics in *(Appx. 5)* shows the differences in average scores between male and female students for the evaluation (modal qualifiers) indicator. Male students have an average value of 1.73, and female students have an average value of 2.37. This demonstrates that female students have the highest average score. *As a result, female students outperformed male students on the evaluation (modal qualifiers) indicator.*

In addition, the descriptive statistic in *(Appx. 5)* shows the differences in average scores between male and female students for the self-regulation (rebuttals) indicator. Male students have an average value of 2.13, and female students have an average value of 2.07. This demonstrates that male students have the highest average score. *As a result, male students outperformed female students on the self-regulation (rebuttals) indicator.*

Table 4.2: Descriptive Statistics of Critical Thinking

	Mean		**Median**		**Std. Deviation**	
	Male	Female	Male	Female	Male	Female
Critical Thinking	13.533	13.733	12	13	4.853	6.606

Table 4.2 explains the differences in average scores between male and female students for the critical thinking. Male students have an average value of 13.53, and female students have an average value of 13.73. This demonstrates that female students have the highest average score. *As a result, female students outperformed male students on the critical thinking.*

Based on the descriptive statistics in *(Appx. 5)* and table 4.2, the researcher came to the tentative conclusion that male students did better than female students on all four indicators of critical thinking: the interpretation (claims) indicator, the

analysis (grounds) indicator, the inference (warrants) indicator, and the self-regulation (rebuttals) indicator. Meanwhile, female students outperformed male students on two indicators of critical thinking: the explanation (backings) indicator and the evaluation (modal qualifiers) indicator. In addition, female students outperformed male students on critical thinking. However, the average (mean) score does not support this statement as a definitive conclusion. Thus, to strengthen the differences between male and female students on the six indicators of critical thinking and the critical thinking score, the researcher analyzes the instrument's scores using the Mann-Whitney test, and the results are as follows:

Table 4.3: Mann-Whitney Test

Critical Thinking Indicator	**Mann-Whitney Test**	
	Mann-Whitney U Score	Asymp. Sig. (2-tailed)
Interpretation (Claims)	96	0.474
Analysis (Grounds)	104	0.714
Inference (Warrants)	102	0.637
Explanation (Backings)	92.5	0.388
Evaluation (Modal Qualifiers)	79.5	0.149
Self-regulation (Rebuttals)	105	0.745
Critical Thinking	110.5	0.934

Table 4.3 shows whether there is a significant difference between male students and female students on the six indicators of critical thinking and the critical thinking score. Then, the researcher explains the results of the Mann-Whitney test in the following order: the interpretation (claims) indicator, the analysis (grounds) indicator, the inference (warrants) indicator, the explanation (backings) indicator, the evaluation (modal qualifiers) indicator, the self-regulation (rebuttals) indicator, and the critical thinking score. The Sig. (2-tailed) column displays the results of the Mann-Whitney test. The results of the Sig. (2-tailed) have the decision criteria of the Mann-Whitney test, which states that if the value of Sig. (2-tailed) < 0.05, then

H0 is rejected and Ha is accepted, which means there is a significant difference between male and female students in critical thinking. Meanwhile, if the value of Sig. (2-tailed) > 0.05, then H0 is accepted and Ha is rejected, which means there is no significant difference between male and female students in critical thinking. Then, the value of the Sig. (2-tailed) for the interpretation (claims) indicator is 0.474, which is higher than 0.05. *As a result, there is no significant difference between male and female students for the interpretation (claims) indicator.*

The results of the Mann-Whitney test for analysis (grounds) indicator can be read in the Sig. (2-tailed) column. It can be seen *(Table 4.3/p. 34)* that the Sig. (2-tailed) value is 0.714, which higher than 0.05. *As a result, there is no significant difference between male and female students in terms of the analysis (grounds) indicator.*

Then, the result of the Mann-Whitney test for the inference (warrants) indicator can be read in the Sig (2-tailed) column. It can be seen *(Table 4.3/p.34)* that the Sig. (2-tailed) value is 0.637, which higher than 0.05. *As a result, there is no significant difference between male and female students in terms of the inference (warrants) indicator.*

The results of the Mann-Whitney test for explanation (backings) indicator can be read in the Sig. (2-tailed) column. It can be seen *(Table 4.3/p.34)* that the Sig. (2-tailed) value is 0.388, which higher than 0.05. *As a result, there is no significant difference between male and female students in term of the explanation (backings) indicator.*

Moreover, the results of the Mann-Whitney test for evaluation indicator (modal qualifiers) can be read in the Sig (2-tailed) column. It can be seen *(Table 4.3/p. 34)* that the Sig. (2-tailed) value is 0.149, which higher than 0.05. *As a result, there is no significant difference between male and female students in term of the evaluation (modal qualifiers) indicator.*

The results of the Mann-Whitney test for self-regulation (rebuttals) indicator can be read in the Sig. (2-tailed) column. It can be seen *(Table 4.3/p. 34)* that the Sig. (2-tailed) value is 0.745, which higher than 0.05. *As a result, there is no significant difference between male and female students in term of the self-regulation (rebuttals) indicator.*

In addition, the results of the Mann-Whitney test for the critical thinking score can be read in the Sig. (2-tailed) column. It can be seen *(Table 4.3/p. 34)* that the Sig. (2-tailed) value is 0.934, which higher than 0.05. *As a result, there is no significant difference between male and female students for the critical thinking.*

Based on the Mann-Whitney test of six critical thinking indicators and the critical thinking score, the researcher concludes that there are no significant differences between male and female students. Despite the fact that the average score (mean) describes the difference in scores between male and female students, it is visible and does not allow for the drawing of any definitive conclusions due to slightly different scores.

4.2 Discussion

Critical thinking is a crucial skill required in the 21st century, particularly in language education. Critical thinking is the process of comprehending the arguments that surround it, as well as the thought processes of students in order to act or take action (Duran & Dokme, 2016; Sumarna et al., 2017; Marfu'i et al., 2019; Sinurat et al., 2020). It is essential to recognize that each individual, regardless of gender, demonstrates a distinct thinking style. Recent literature has shown gender differences in the use of writing strategies and critical thinking abilities among students (Bai, Shen, & Mei, 2020; Liu & Pasztor, 2023). Thus, in this study, the researcher combined the indicators of critical thinking (Facione, 2015) with the indicators of argumentative writing (Toulmin et al., 2002) to score the students' argumentative writing and determine the significant differences between male and female students in critical thinking through argumentative writing.

Furthermore, in this study, the first indicator of critical thinking is interpretation (claims). Interpretation is the knowledge of how to convey the importance or meaning of several kinds of experiences, situations, data, events, assessments, habits, customs, beliefs, rules, procedures, or criteria (Seventika et al., 2018; Facione, 2015). It was in line with the definition of claims, which refers to statements that support, clarify, express, reject, or demand anything that relies on an objective argument or statements (Magalhaes, 2020; Toulmin et al., 2002). Then, this study's findings revealed that male students were superior to female students in terms of average interpretation (claims). The findings from descriptive statistics *(Table 4.3/p. 33)* indicate that male students have an average value of 2.67, while female students have an average value of 2.27. However, there is no significant difference between male and female students for the interpretation (claims) indicator.

The second indicator of critical thinking is analysis (grounds). Analysis is the ability to identify between the statements, the questions, and the information intended to express or support the first statement (Marni et al., 2020; Seventika et al., 2018; Facione, 2015). It has the same meaning as grounds, which refers to the foundation with the existing facts and data to support claims (Magalhaes, 2020; Toulmin et al., 2002). This study's findings revealed that male students were superior to female students in terms of average analysis (grounds). The findings from descriptive statistics *(Table 4.3/p. 33)* indicate that male students have an average value of 2.33, while female students have an average value of 2.27. However, there is no significant difference between male and female students for the analysis (grounds) indicator.

In addition, inference (warrants) is the third critical thinking indicator. Inference is the ability to identify the statements as more logically concluded and reasonable judgments (Marni et al., 2020; Seventika et al., 2018; Facione, 2015), and warrants are the absolute statements that are connected to and supported by the claims and grounds (Magalhaes, 2020; Toulmin et al., 2002). Furthermore, this study's findings revealed that male students were superior to female students in

terms of average inference (warrants). The findings from descriptive statistics *(Table 4.3/p. 33)* indicate that male students have an average value of 2.40, while female students have an average value of 2.27. However, there is no significant difference between male and female students for the inference (warrants) indicator.

Then, the fourth critical thinking indicator is explanation (backings). Explanation is the ability to establish the reason for the statement more logically based on the existing data (Seventika et al., 2018; Facione, 2015), and backing refers to the ability to extend the statement to be true and valid based on the data (Magalhaes, 2020; Toulmin et al., 2002). Furthermore, this study's findings revealed that female students were superior to male students in terms of average explanations (backings). The findings from descriptive statistics *(Table 4.3/p. 33)* indicate that male students have an average value of 2.27, while female students have an average value of 2.60. However, there is no significant difference between male and female students in terms of the explanation (backings) indicator.

Furthermore, evaluation (modal qualifiers) is the fifth critical thinking indicator. Evaluation is the process of looking at or questioning existing data and figuring out what it means with the help of a logical statement (Seventika et al., 2018; Facione, 2015). Modal qualifiers are used to make the conclusion stronger so that the statement is valid (Magalhaes, 2020; Toulmin et al., 2002). This study's findings revealed that female students were superior to male students in terms of average evaluation (modal qualifiers). The findings from descriptive statistics *(Table 4.3/p. 33)* indicate that male students have an average value of 1.73, while female students have an average value of 2.27. However, there is no significant difference between male and female students for the evaluation (modal qualifiers) indicator.

Moreover, the sixth critical thinking indicator is self-regulation (rebuttals). Seventika et al., 2018; Facione, 2015 define self-regulation as the ability to understand and control one's behavior and responses to emotions, while Magalhaes, 2020; Toulmin et al., 2002 define rebuttals as exceptions to the claim. This study's findings revealed that male students were superior to female students in terms of

average inference (warrants). The findings from descriptive statistics *(Table 4.3/p. 33)* indicate that male students have an average value of 2.13, while female students have an average value of 2.06. However, there is no significant difference between male and female students for the self-regulation indicator. The six critical thinking indicators above demonstrate that there is no significant difference between male and female students. The research (Marni et al., 2020) supports the findings of this study, revealing no significant differences between male and female students in each indicator. However, this study used six indicators with no significant difference between male and female students, whereas Marni et al. (2020) used four indicators, such as interpretation, analysis, inference, and evaluation, and examined the eight sub-indicators, which revealed seven indicators had no significant differences and one indicator had a significant difference.

Furthermore, this study's critical thinking average score revealed that female students were superior to male students in critical thinking. However, several studies have demonstrated that the critical thinking skills of male students are superior to those of female students (Andayani et al., 2020; Azar, 2010; Liu et al., 2019; Piaw, 2014; Preiss et al., 2013; Verawati et al., 2010). Then, previous studies did not assess the indicators of critical thinking. This study examined the students' critical thinking through their argumentative essay, using six indicators of critical thinking. This study assessed the six indicators of critical thinking, namely interpretation, analysis, inference, explanation, evaluation, and self-regulation (Facione, 2015). The results of this study revealed that male students were superior in four indicators of critical thinking compared to female students, and female students were superior in two indicators of critical thinking, even though there were no significant differences between male and female students in critical thinking. The research (Kawuryan & Aman, 2022; Budi, 2017; Hayudiyanti et al., 2017; Azin & Tabrizi, 2016) supports these findings, indicating no significant differences in critical thinking between male and female students. However, the lack of raters' criteria in this study stems from the fact that the raters originate from different universities than the students. To mitigate this issue, the researcher proposes that future researchers involve raters from the same university as the students. The

researcher can rely on the level of percent agreement to determine the inter-rater reliability test, provided that the raters are well-trained and there is minimal chance of guesswork in their scoring (McHugh, 2012).

CHAPTER V

CONCLUSION AND RECOMMENDATION

This chapter convey a conclusion and recommendation. The conclusion will conclude the research findings, then the recommendation will be offered to several parties.

5.1 Conclusion

The findings of the study concluded that there are no significant differences between male and female students in critical thinking skills. This study synthesized critical thinking and argumentative indicators to measure the differences between male and female students in critical thinking. Thus, a causal comparative study, or ex pos facto, was used to determine the differences, and the sample size was small. This study involved 15 male students and 15 female students, and a nonparametric statistic, the Mann-Whitney test, was used as the differentiate test to determine male and female student differences in critical thinking.

Furthermore, the six critical thinking indicators indicated that there were no differences between male and female students. However, when viewed as the average or mean score of the six indicators, male students were superior to female students in four indicators: interpretation, analysis, inference, and self-regulation. Then, female students outperformed male students in two indicators: explanation and evaluation. This may be due to the fact that the study used two raters who gave scores that were not slightly different. However, in the total score of indicators that indicate the critical thinking score, female students were superior to male students in critical thinking.

5.2 Recommendation

According to the study's results, the researcher offered several recommendations for some parties: the institution and further research.

The Institution

The study can be a reference for teachers and university lectures to determine students' critical thinking. The study can be a reference to the material to include critical thinking, such as theories or indicators for the curriculum, which can help teachers, lecturers, and even students to be more conscious of critical thinking.

Further Research

This critical thinking study should be taught to secondary or high school students in terms of developing critical thinking abilities. Moreover, the technique of sampling with the small sample can be taken into consideration by the researcher for future research involving more than one university student or the total number of students' essays on a large population. Then, the next researcher highly recommends using a mixed-methods research design, where the quantitative descriptive method measures and reveals the scores of students' essays, and the qualitative method interprets the results. Furthermore, the next researcher may use alternative scoring and testing methods to measure critical thinking ability.

REFERENCES

Al-Saadi, Z. (2020). Gender differences in writing: the mediating effect of language proficiency and writing fluency in text quality. *Cogent Education, 7*(1). https://doi.org/10.1080/2331186X.2020.1770923

Andayani, S., Muchtar, H., Yufiarti., & Susanto, E. (2020). The profile of critical thinking and learning outcomes of teacher candidates viewed from gender differences. *Journal of Physics: International Conference on Mathematics, Science, and Education, 1567.* https://doi.org/10.1088/1742-6596/1567/4/042072

Andrews, R., Freeman, A., Hou, D., McGuinn, N., Robinson, A., & Zhu, J. (2018). The effects of argumentation writing and source-based writing on the development of students' argumentation skills and source-based writing skills. *Journal of Writing Research, 9*(3), 277-315.

Archer, J., & Lloyd, B. (2002). *Sex and gender* (2nd ed.). Cambridge University Press.

Ary, D., Jacobs, C. L., Sorensen, C. K., & Walker, D. A. (2014). *Introduction to research in education* (9th ed.). Wadsworth Cengage Learning.

Asterhan, C. S. C. (2018). Exploring enablers and inhibitors of productive peer argumentation: The role of individual achievement goals and of gender. *Contemporary Educational Psychology, 54*(1), 66-78. https://doi.org/10.1016/j.cedpsych.2018.05.002

Asterhan, C. S. C., Schwarz, B. B., & Gill, J. (2012). Small-group, computer-mediated argumentation in middle-school classrooms: The effects of gender and different types of online teacher guidance. *British Journal of Educational Psychology, 82*(3), 375-397. https://doi.org/20.2222/j.2044-8279.2011.02030.x

Bai, B., Shen, B., & Mei, H. (2020). Hong Kong primary students' self-regulated writing strategy use: influence of gender, writing proficiency, and grade level. *Studies in Educational Evaluation, 65,* 1-11. https://doi.org/10.1016/j.stucduc.2020.100839

Bassham, G., Irwin, W., Nardone, H., & Wallace, James W. (2018). *Critical Thinking: A Students's Introduction* (6th ed.). McGraw-Hill.

Bazerman, C. (2016). What to sociocultural studies of writing tell us about learning to write? In C.A. MacArthur, S. Graham, & J. Fitzgerald (Eds,), *Handbook of writing research* (2nd ed.), 11-23. Guilford.

Birkenstein, C., & Graff, G. (2018). *They say/I say* (4th Ed.). W. W. Norton & Company.

Cáceres, M., Nussbaum, M., & Ortiz, J. (2020). Integrating critical thinking into the classroom: A teacher's perspective. *Thinking Skills and Creativity, 37*, 100674. https://doi.org/10.1016/j.tsc.2020.100674

Cohen, J. (1968). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 32-46. https://doi.org/10.1177/001316446002000104.

Conover, W. J. (1999). *Practical nonparametric statistics.* John Wiley & Sons.

Cottrell, S. (2023). *Critical thinking skills: Effective analysis, argument and reflection* (4th ed.). Bloomsbury Publishing.

Creswell, J. W. (2018). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (5th ed.) Pearson Education, Inc.

Darmaji, D., Astalini, A., Kurniawan, D. A., & Br. Ginting, A. A. (2021). Relationship of science process skills on critical thinking ability review by gender in madrasah aliyah. *Jurnal Pendidikan Sains Indonesia, 9*(4), 711–735. https://doi.org/10.24815/jpsi.v9i4.21391

Douglas, D. (2010). *Understanding language testing.* Routledge.

Duran, M., & Dokme, I. (2016). The effect of the inquiry-based learning approach on students' critical-thinking. *Eurasia Journal of Mathematics, Science & Technology Education, 12*(12), 2887-2908. https://doi.org/10.12973/Eurasia.2016.02311a

Edmonds, W. A., & Kennedy, T. D. (2017). *An applied guide to research designs* (2nd ed). SAGE.

Ennis, R. H., & Norris, S. (1989). Evaluating critical thinking. *The Practitioners' Guide to Teaching Thinking Series.* Pacific Grove, Midwest Publication.

Ennis, R. (2011). Critical thinking: reflection and perspective part ii. *Inquiry: Critical Thinking Across the Disciplines, 26*(2), 5-19. https://doi.org/10.5840/inquiryctnews201126215

Facione, P. A., & Gittens C. A. (2016). *Think Critically*. Pearson Education, Inc.

Ferretti, R. P., & Graham, S. (2019). Argumentative writing: theory, assessment, and instruction. *Reading and Writing, 32*, 1345-1357. https://doi.org/10.1007/s11145-019-09950-x

Ferretti, R. P., & Fan, Y. (2016). Argumentative writing. *Best practices in writing instruction, 135*. In C. A. MacArthur., S. Graham., & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed.), pp. 301-315. Guilford Press.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to evaluate research in education* (8th ed.). McGraw-Hill

Garnison. D. R., Anderson, T. & Archer, W. (2001). Critical thinking and computer conferencing: a model and tool to assess cognitive presence. *American Journal of Distance Education 15*(1), 7-23. http://hdl.handle.net/2149/740

Golombok, S., Rust, J., Zervoulis, K., Croudace, T., Golding, J., & Hine, M. (2008). Developmental trajectories of sex-typed behaviour in boys and girls: a longitudinal general population study of children aged 2.5-8 years. *Child Development, 79*(5), *1583-1593.* https://doi.org/10.1111/j.1467-8624.2008.01207.x

Graham, S. (2018). The writers-within-community model of writing. *Educational Psychologist. 53*(4), 258-279. https://doi.org/10.1080/00461520.2018.1481406

Graham, S., Harris, K., & Hebert, M. (2011). *Informing writing: The benefits of formative assessment.* Carnegie Corporation.

Heaton, J. B. (1988). *Writing English Language Tests.* Longman, Inc.

Huang, G. C., Lindell, D., Jaffe, L. E., & Sullivan, A. M. (2016). A multi-site study of strategies to teach critical thinking: "Why do you think that?" *Medical Education*, *50*(2), 236–249. https://doi.org/10.1111/medu.12937

Hyland, K. (2013). *Materials for developing writing skills*. Bloomsbury Academic.

Ichsanda, B., & Hz, R. (2022). Exploring gender differences in public speaking nxiety. *Sawwa: Jurnal Studi Gender 17*(2), 247–266. https://doi.org/10.21580/sa.v17i2.14295

Islam, K. M. B. (2018). Issues in language teaching: gender in language education. *Intellection: A Bi-annual Interdisciplinary Research Journal. 6*(2), 1-8.

Kawuryan, S. P., Sayuti, S. A., & Aman. (2022). Critical thinking among fourth grade elementary school students: A gender perspective. *Cakrawala Pendidikan: Jurnal Ilmiah Pendidikan, Vol. 41*(1), 211-224. https://doi.org/10.21831/cp.v41i1.44322

Kim, M. K., & Pollard, V. A. (2017). A modest critical pedagogy for English as a foreign language education. *Education as Change*, *21*(1), 50–72. https://doi.org/10.17159/1947-9417/2017/492

Keynes, M. (2008). *Thinking critically.* United Kingdom: Thanet Press.

Kothari, C. R. (2004). *Research methodology: methods & techniques.* New Age International Ltd.

Kuhn, D. (1991). *The skills of argument.* Cambridge University Press.

Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174. https://doi.org/10.2307/2529310

Lai, E. R. (2011). Critical thinking: A literature review. *Pearson's Research Reports*, *6*(1), 40-41.

Lin, Y. (2018). Developing critical thinking in EFL classes. *Springer Singapore*, *1*, *194.* https://doi.org/10.1007/978-981-10-7784-5

Liu, Y., & Pasztor, A. (2023). Moderated mediating effects of gender among the components of critical thinking disposition in undergraduate students. *Heliyon, 9.* https://doi.org/10.1016/j.heliyon.2023.el4664

Magalhaes, A. L. (2020). Teaching how to develop an argument using the toulmin model. *International Journal of Multidisciplinary and Current Educational Research, 2*(2), 01-07.

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics, 18*(1), 50-60. https://doi.org/10.1214/aoms/1177730491

Marfui, L. N. R., Ilfiandra, & Nurhudaya. (2019). The analysis of critical thinking skills test in social-problems for physics education students with rasch model. *Journal of Physics: Conference Series, 1280(5).* https://doi.org/10.1088/1742-6596/1280/5/052012

Marni, S., Aliman, M., Suyono, Roekhan, & Harsiati, T. (2020). Students' critical thinking skills based on gender and knowledge group. *Journal of Turkish Science Education, 17*(4), 544-560. https://www.tused.org/index.php/tused/article/view/1246

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica, 22*(3), 276-82. https://hrcak.srce.hr/89395

Mertler, C. A. (2022). *Introduction to educational research* (3rd ed). SAGE Publication.

Moon, J. (2008). *Critical Thinking: An Exploration of Theory and Practice.* Routledge.

Moose, J. C. (2003). *Gender dan pembangunan alih bahasa Hartian Silawati.* Pustaka Pelajar.

Newell, G., Bloome, D., Kim, M. Y., & Goff, B. (2018). Shifting epistemologies during instructional conversations about "good" argumentative writing in a high school English language arts classroom. *Reading and Writing: An Interdisciplinary Journal, 32,* 1359-1382. https://doi.org/10.1007/s11145-018-9905-y.

Noroozi, Omid., Hatami, Javad., Bayat, Arash., Ginkel, Stan van., Biemans, Harm J.A., & Mulder, Martin. (2018): Students' online argumentative peer feedback, essay writing, and content learning: Does gender matter? *Interactive Learning Environments, 28*(6), 698-712. https://10.1080/10494820.2018.1543200

Noroozi, O., Banihashem, S. K., Kerman, N. T., Khaneh, M. P. A., Babayi, M., Ashrafi, H., & Biemans, H. J.A. (2022): Gender differences in students' argumentative essay writing, peer review performance and uptake in online learning environments. *Interactive Learning Environments, 31*(10), 6302-6316. https://10.1080/10494820.2022.2034887

Noroozi, O., Kirschner, P. A., Biemans, H. J. A., & Mulder, M. (2018). Promoting argumentation competence: Extending from first- to second-order scaffolding through adaptive fading. *Educational Psychology Review, 30*, 153–176. https://doi.org/10.1007/s10648-017-9400-z

Noroozi, O., Biemans, H. J. A., & Mulder, M. (2016). Relations between scripted online peer feedback processes and quality of written argumentative essay. *Internet and Higher Education, 31*(1), 20–31. https://doi.org/10.1016/j.iheduc.2016.05.002

Nussbaum, E. M., & Schraw, G. (2007). Promoting argument–counterargument integration in students' writing. *The Journal of Experimental Education, 76*(1), 59–92. https://doi.org/10.3200/JEXE.76.1.59-92

O'Keefe, D. J. (1999). How to handle opposing arguments in persuasive messages: A meta-analytic review of the effects of one-sided and two-sided messages. *Annals of the International Communication Association*, *22*(1), 209-249. https://doi.org/10.1080/23808985.1999.11678963

Pahamzah, J., Andriani, H., Gailea, N., & Baihaqi, A. (2022). The correlation among plagiarism awareness, genders, and essay writing skills of English education of postgraduate students of Sultan Ageng Tirtayasa university. *Litera*, *21*(1), 1–8. https://doi.org/10.21831/ltr.v21i1.40649

Paul, R., & Elder, L. (2009*). The Miniature Guide to Critical Thinking: Concepts and Tools* (6th ed.). Kindle Edition.

Penkauskiene, D., Railiene, A., & Cruz, G. (2019). How is critical thinking valued by labour market? Employer perspectives from different European countries. *Studies in Higher Education, 44(5), 804-815.* https://doi.org/10.1080/03075079.2019.1586323

Preiss, D. D., Castillo, J. C., Flotts, P., & Martin, E. S. (2013). Assessment of argumentative writing and critical thinking in higher education: Educational correlates and gender differences. *Learning and Individual Differences, 28,* 193-203. http://dx.doi.org/10.1016/j.lindif.2013.06.004

Ravid, R. (2020). *Practical Statistics for Educators.* Rowman & Littlefield.

Rianto, A. (2021). Examining gender differences in reading strategies, reading skills, and English proficiency of EFL University students. *Cogent Education*, *8*(1), 1993531. https://doi.org/10.1080/2331186X.2021.1993531

Seventika, S. Y., Sukestiyarno, Y. L., & Mariani, S. (2018). Critical thinking analysis based on Facione (2015) – Angelo (1995) logical mathematics

material of vocational high school (VHS). *International Conference on Mathematics, Science, and Education, 983*(1). https://doi.10.1088/1742-6596/983/1/012067

Setyowati, L., Agustina, F., Sukmawan, S., El-Sulukiyyah, A. A., & Mabaroh, B. (2020). The students' problems and solutions in writing an argumentative essay on gender issue. *Journal of English Language Teaching and Linguistics, 5*(3), 279-292. www.jeltl.org

Sinurat, R., Nevrita, N., & Hindrasti, N.E.K. (2020). Identifikasi tingkat kemampuan berpikir kritis siswa pada materi asi eksklusif dan program keluarga berencana. *Jurnal IPA & Pembelajaran, 4(1):60-69.* https://doi.org/10.24815/Jipi.V4i1.15728

Soodmand Afshar, H., Movassagh, H., & Radi Arbabi, H. (2017). The interrelationship among critical thinking, writing an argumentative essay in an L2 and their subskills. *The Language Learning Journal, 45(5), 419-433.* https://doi.org/10.1080/09571736.2017.1320420

Straková, Z., & Cimermanová, I. (2018). Critical thinking development-a necessary step in higher education transformation towards sustainability. *Sustainability (Switzerland), 10*(10). https://doi.org/10.3390/su10103366

Steinkamp, M. W., & Maehr, M. L. (1983). Affect, ability, and science achievement: a quantitative synthesis of correlational research. *Review of Educational Research, 53*(3), 369-396. https://doi.org/10.3102/00346543053003369

Sumarna, N., Wahyudin, & Herman, T. (2017). Preface: International conference on recent trends in physics (Icrtp 2016). *Journal of Physics: Conference Series, 755*(1), 1-9. https://doi.org/10.1088/1742-6596/755/1/011001

Talbot, M. (2010). *Language and gender* (2nd ed.). MPG Books Group: Great Britain.

Tilaar, H. A. R. (2011). *Pedagogik kritis: Critical pedagogy*. Rineka Cipta.

Toulmin, S. E., Rieke, R., & Janik, A. (2002). *An Introduction to Reasoning* (2nd ed.). Mac-millan Publishing.

Tsemach, E., & Zohar, A. (2021). The intersection of gender and culture in argumentative writing. *International Journal of Science Education, 43*(6), 969–990. https://doi.org/10.1080/09500693.2021.1894499

UNESCO. (2017). Education for Sustainable Development Goals (SDGs). In *European Conference on Educational Research 2017*.

van Eemeren, F. H. (2018). *Argumentation theory: A pragma-dialectical perspective.* Springer, Cham.

Wentzel, A. (2017). *A guide to argumentative research writing and thinking: overcoming challenges* (1st ed.). Routledge.

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

Wu, S. M. (2006). Creating a contrastive rhetorical stance: Investigating the strategy of problematization in students' argumentation. *Regional Language Centre Journal, 37*(3), 329–353. https://doi.org/10.1177/0033688206071316

APPENDICES

Appendix 1. Indicators of Critical Thinking for Scoring Students' Argumentative Writing

No.	Indicators	Indicator's Description	Score
1.	Interpretation (Claims)	The student cannot write and express opinions and the meaning of certain issues; many errors in word choice, diction, and grammar distract or confuse the reader.	0
		The student writes and expresses opinions and the meaning of certain issues but is flawed in word choice, diction, and grammar, and they distract or confuse the reader.	1
		The student writes and expresses opinions and the meaning of certain issues, as well as frequent errors in word choice, diction, and grammar, which distract or confuse the reader.	2
		The student writes and expresses opinions and the meaning of certain issues correctly but not comprehensively, with occasional errors in word choice, diction and grammar, which do not distract or confuse the reader.	3
		The student writes and expresses opinions and the meaning of certain issues correctly and comprehensively, effectively using word choice, diction and grammar, which the reader could outline.	4
2.	Analysis (Grounds)	The student cannot identify and analyse the issues based on the existing truth.	0
		Lack of competence to identify and analyse the issues based on the existing truth.	1
		The student is adequate to identify and analyse the issues based on the existing truth.	2
		The student identifies and analyses the issues correctly but not comprehensively based on the existing truth.	3
		The student identifies and analyses the issues correctly and comprehensively based on the existing truth.	4
3.	Inference (Warrants)	The student cannot conclude, consider information, and review an argument according to the existing problem based on facts and data.	0
		Lack of ability to conclude, consider information, and review an argument according to the existing problem based on facts and data.	1
		The student concludes but does not consider information or review an argument according to the existing problem based on facts and data.	2

		The student concludes, considers information, or reviews an argument correctly but not comprehensively according to the existing problem based on facts and data.	3
		The student concludes, considers information, or reviews an argument correctly and comprehensively according to the existing problem based on facts and data.	4
4.	**Explanation (Backings)**	The student cannot conclude certain issues that have been explained and does not present them in an argument that convinces the reader.	0
		There is a lack of conclusion about certain issues that have been explained and does not present them in an argument that convinces the reader.	1
		The student is adequate to conclude certain issues that have been explained and does not present them in an argument that convinces the reader.	2
		The student concludes certain issues that have been explained and presents them in an argument correctly but not comprehensively to convince the reader.	3
		The student concludes certain issues that have been explained and presents them in an argument correctly and comprehensively to convince the reader.	4
5.	**Evaluation (Modal Qualifiers)**	The student cannot evaluate the arguments and statements of certain issues.	0
		Lack of competence to evaluate arguments and statements of certain issues in the text.	1
		The student is adequate in evaluating arguments and statements of certain issues in the text but may have flaws that make it difficult to convince the reader	2
		The student is competent in evaluating arguments and statements of certain issues in the text and can convince the reader correctly but not comprehensively.	3
		The student is competent in evaluating arguments and statements of certain issues in the text and can convince the reader correctly and comprehensively.	4
6.	**Self-regulation (Rebuttals)**	Student cannot evaluate and analyze certain issues according to their cognition, self-awareness, and application in life.	0
		Lack of competence to evaluate and analyze certain issues in their self-awareness, but not according to their cognition and application in life.	1

		Lack of competence to evaluate and analyze certain issues in their self-awareness and cognition, but not according to application in life.	2
		The student is adequately competent in evaluating and analyzing certain issues according to their cognition, self-awareness, and application in life.	3
		The student is competent in evaluating and analyzing certain issues comprehensively according to their cognition, self-awareness, and application in life.	4
Total Score			24

{Indicators of Critical Thinking (Facione, 2015) and Indicators of Argumentative Writing (Heaton, 1988; Toulmin et al., 2002)}

Appendix 2. The Critical Thinking Score of The First-rater

No.	Gender	IC	AG	IW	EB	EMQ	SrR	CT Score
1.	M	1	1	2	1	1	1	7
2.	M	1	1	2	1	1	1	7
3.	M	1	1	2	1	1	1	7
4.	F	1	1	1	1	1	1	6
5.	F	1	1	1	1	1	1	6
6.	F	1	1	1	1	1	1	6
7.	M	3	2	2	2	1	2	12
8.	M	3	2	2	2	1	2	12
9.	M	3	4	2	2	3	2	16
10.	M	3	3	3	3	2	2	16
11.	F	1	1	2	2	2	1	9
12.	F	1	1	2	2	2	1	9
13.	F	1	1	2	2	2	1	9
14.	F	2	3	3	4	3	3	18
15.	F	4	4	3	3	2	2	18
16.	F	3	4	2	3	3	3	18
17.	M	3	2	2	2	2	3	14
18.	F	4	3	3	4	4	3	21
19.	M	3	3	4	3	3	3	19
20.	F	4	4	4	4	3	4	23
21.	F	4	3	3	3	4	3	20
22.	F	4	4	4	4	3	4	23
23.	M	2	2	3	3	2	2	14
24.	M	3	2	2	2	1	1	11
25.	M	4	4	3	4	3	4	22
26.	M	4	4	3	4	3	4	22
27.	F	1	1	1	2	1	1	7

28.	F	2	2	2	3	2	2	13
29.	M	3	2	2	2	1	2	12
30.	M	3	2	2	2	1	2	12

Note: **M:** Male; **F:** Female

IC: Interpretation (Claims)

AG: Analysis (Grounds)

IW: Inference (Warrants)

EB: Explanation (Backings)

EMQ: Evaluation (Modal Qualifiers)

SrR: Self-regulation (Rebuttals)

CT: Critical Thinking

Appendix 3. The Critical Thinking Score of The Second-rater

No.	Gender	IC	AG	IW	EB	EMQ	SrR	CT Score
1.	M	1	1	2	1	2	1	8
2.	M	1	1	2	1	2	1	8
3.	M	1	1	2	1	2	1	8
4.	F	1	2	2	1	1	1	8
5.	F	1	2	2	1	1	1	8
6.	F	1	2	2	1	1	1	8
7.	M	3	2	2	3	2	3	15
8.	M	3	2	2	3	2	3	15
9.	M	2	4	2	3	3	2	16
10.	M	3	4	3	3	3	3	19
11.	F	2	1	2	2	2	1	10
12.	F	2	1	2	2	2	1	10
13.	F	2	1	2	2	2	1	10
14.	F	3	3	3	4	3	2	18
15.	F	4	4	3	3	3	3	20
16.	F	3	4	3	3	3	3	19
17.	M	4	2	2	2	3	3	16
18.	F	4	3	3	4	4	3	21
19.	M	4	3	4	3	3	4	21
20.	F	4	4	4	4	3	3	22
21.	F	4	3	3	4	4	4	22
22.	F	4	4	3	4	4	4	23
23.	M	3	3	3	3	2	2	16
24.	M	3	3	2	2	2	2	14
25.	M	4	4	3	4	3	3	21
26.	M	4	4	3	4	3	3	21
27.	F	2	2	2	3	2	2	13

28.	F	3	3	2	3	3	3	17
29.	M	3	2	2	3	2	3	15
30.	M	3	2	2	3	2	3	15

Note: **M:** Male; **F:** Female

IC: Interpretation (Claims)

AG: Analysis (Grounds)

IW: Inference (Warrants)

EB: Explanation (Backings)

EMQ: Evaluation (Modal Qualifiers)

SrR: Self-regulation (Rebuttals)

CT: Critical Thinking

Appendix 4. Normality Test

	Shapiro-Wilk Normality Test			
Indicators	**Gender**	**Statistic**	**df**	**Sig.**
Interpretation (Claims)	Male	.779	15	.002
	Female	.743	15	.001
Analysis (Grounds)	Male	.849	15	.017
	Female	.770	15	.002
Inference (Warrants)	Male	.667	15	.000
	Female	.885	15	.056
Explanation (Backings)	Male	.872	15	.037
	Female	.876	15	.042
Evaluation (Modal Qualifiers)	Male	.734	15	.001
	Female	.885	15	.056
Self-regulation (Rebuttals)	Male	.847	15	.016
	Female	.804	15	.004
Critical Thinking	Male	.920	15	.195
	Female	.863	15	.027

Appendix 5. Descriptive Statistics from Six Indicator of Critical Thinking

Indicators	Mean		Median		Std. Deviation	
	Male	Female	Male	Female	Male	Female
Interpretation (Claims)	2.667	2.267	3	2	0.976	1.387
Analysis (Grounds)	2.333	2.267	2	2	1.047	1.335
Inference (Warrants)	2.400	2.267	2	2	0.632	1.033
Explanation (Backings)	2.267	2.600	2	3	0.961	1.121
Evaluation (Modal Qualifiers)	1.733	2.267	1	2	0.884	1.033
Self-regulation (Rebuttals)	2.133	2.067	2	2	0.990	1.163

Appendix 6. Students' Essay with Highest Score

Female Student

Total Score: 23

Aerobic Fitness

Aerobic fitness can be defined as the ability of the body's cardiovascular and muscular systems to provide the necessary energy to sustain activity that uses the large muscle groups over an extended period of time (Interpretation/Claims). To reach aerobic fitness, a person must engage in continuous activity like jogging, walking, cycling, and stair climbing, rowing, or swimming at an intensity level you can maintain for at least 30 minutes, three to seven days per week (Bruce Rife, 2003) (Analysis/Grounds).

Although considerable epidemiologic and clinical evidence suggests that structured exercise, increased lifestyle activity, or both are cardio protective, the absolute and relative risk of cardiovascular and musculoskeletal complications appear to increase transiently during vigorous physical activity (Analysis/Grounds).

According to Henry Winston (2006), You may be wondering what the benefits of aerobic exercise actually are. Sure you hear everywhere that performing regular exercise is good for your health, but why is that? The answer is actually quite long because there is a wide array of advantages to be had from doing some kind of aerobic exercise on a daily basis. It helps keep your heart, brain, lungs, and your body healthy, plus a whole lot more to **(**Interpretation/Claims**) thus, aerobic exercise is helpful for regulates weight and mental benefit** (Inference/Warrants**).**

Aerobic exercise alone may hold the power to help you lose weight and keep it off. Depending on your weight and speed, you may need to walk or jog up to 4 miles to burn 400 to 600 calories (Explanation/Backings). However, this idea could not be further developed, because losing weight only with aerobic exercise is difficult if it is not matched by other sports (Evaluation/Modal Qualifiers) there are two causes for athletes experiencing overtraining. First, the nervous system that is too tired. Secondly, training is still done when the energy supply from the body has run out. Be careful both of these conditions have a negative effect on fat burning (Evaluation/Modal Qualifiers). When the body runs out of energy supply, it will look for other energy sources to try to maintain its life. At that time, the remaining alternative energy source is protein (muscle) in the body. Excessive doing aerobics, the body begins to release muscle into sugar energy and catabolism occurs. According to Ade Rai (2012), Catabolism is a condition in which muscle mass shrinks due to overtraining. It is at this time that exercise activity becomes counter-productive because it is excessive **(**Interpretation/Claims**)**. When catabolism occurs, the muscle's calorie-burning capacity decreases, because the amount of muscle has decreased reminds us to avoid this condition in order to achieve maximum calorie and fat burning (Evaluation/Modal Qualifiers).

Aerobic exercise is helpful to regulates weight. You may have heard that diet and exercise are the building blocks to weight loss. But aerobic exercise alone may hold the power to help you lose weight and keep it off (Interpretation/Claims). According to Jared E. Albert (1994) researchers asked overweight participants to

keep their diets the same, but to engage in exercise sessions that would burn either 400 to 600 calories, 5 times a week, for 10 months (Analysis/Grounds). The results showed significant weight loss, between 4.3 and 5.7 percent of their starting weights, for both men and women. Most participants walked or jogged on treadmills for the majority of their exercise sessions. If you don't have access to a treadmill, try taking a few brisk walks or jogs a day, such as during your lunch break or before dinner. Depending on your weight and speed, you may need to walk or jog up to 4 miles to burn 400 to 600 calories. Cutting calories in addition to aerobic exercise can reduce the amount of exercise needed to lose the same amount of weight (Inference/Warrants).

Aerobic exercise is also helpful in mental benefits for health. There is a large percentage of the population that suffers from things like anxiety, depression, and even low self-confidence. It was previously thought that all of these things could only be fixed with drugs or psychiatric help, but that does not seem to be the case (Interpretation/Claims). Regular exercise can actually be quite an effective tool when it comes to curing anxiety or depression. Studies have shown that this has something to do with a phenomenon that we call the runner's high. The runner's high is a feeling of elation and happiness that runners get after having exercised for a prolonged period of time. It is shown that aerobic exercise causes the brain to release chemicals called dopamine. Dopamine plus a few other brain chemicals are feel-happy chemicals. When released by the brain they make you feel happy and satisfied. (Karl: 2006) Therefore people who suffer from depression and anxiety can reduce their symptoms through some simple aerobic exercise (Analysis/Grounds). Of course, the fact that exercise helps to boost self-confidence has to do with the fact that you feel fitter and probably look better too (Inference/Warrants).

All in all, (Self-regulation/Rebuttals) an aerobic exercise program is feasible and effective for individuals with traumatic brain injury, which leads to improved mood, cardiovascular fitness, and self-esteem. Besides aerobic exercise can also reduce weight and help in mental benefits for health. Actually, there are many benefits of aerobic exercise, but in my opinion, the two benefits I described above are the most resilient and best benefits for the health of our bodies (Explanation/Backings). Aerobic exercises are having many benefits for our body, so let's do it, so that the body becomes healthier and fitter (Evaluation/Modal Qualifiers). Your health is how you treat it, if you are lazy, it can also be bad for your health (Self-regulation/Rebuttals). So let's exercise regularly with aerobics (Self-regulation/Rebuttals).

Indicators of Critical Thinking for Scoring Argumentative Writing

No.	Indicators of Critical Thinking	Score
1.	Interpretation (Claim)	4
2.	Analysis (Ground)	4
3.	Inference (Warrant)	4
4.	Explanation (Backing)	3
5.	Evaluation (Modal Qualifier)	4
6.	Self-regulation (Rebuttal)	4

Appendix 7. Students' Essay with Middle Score

Male Student

Total Score: 16

Benefits and Risks of New Technology

Definition of new technology is any set of productive techniques which offers a significant improvement (whether measured in terms of increased output or savings in costs) over the established technology for a given process in a specific historical context (Gordon Marshall, 2020).

Almost everything in our lives is becoming reliant on technology, from the food we eat to the work we do (Interpretation/Claims). Even the biological processes such as reproduction in humans, animals, and plants are becoming dependent on new technology. Everyday new technology and innovations are invented and adopted by some people who have access (Interpretation/Claims). Technology has changed many things.

According to Bruce and Brooks (1987), advancements in technology are always intended to "make things easier, save time, and increase efficiency" (Analysis/Grounds). These technologies also come with their costs and downsides, which sometimes are overlooked because of the benefits. Therefore, the new technology is harmful for the society owing to their disruptive influence (Inference/Warrants).

We use technology every day in accomplishing various tasks or interests. It simplifies life, and many people have embraced it for different reasons (Interpretation/Claims). But, it makes people are overly reliant on technology (Inference/Warrants). The more a society advances technologically, the more the people depend on computers and other technological advances in their everyday activities (Interpretation/Claims). As a result, when a machine stops to function or a computer-crashes, people become almost disabled and cannot do anything until the problem is resolved (Inference/Warrants). This kind of reliance on technology, therefore, puts humans at a distinct disadvantage since they become less self-dependent. The new technology also brought advanced communication technology tools that have made it easy to communicate in our daily lives (Interpretation/Claims). But on the other side this makes people spend more time on using social networks, learning how to use new technologies, and playing games hence neglect their real lives. According to Gaby and Kenzie (2013), "new technology makes people lazier, and lack the desire to engage in real life." It made many people less sensitive in the society (Analysis/Grounds). They argue that technology is a privilege to have but it's important to interact with other people (Explanation/Backings).

Firstly, it is clear that modern technology has replaced human labor; robots are undertaking the jobs which were previously done by humans (Ramey Karehka,2012) (Analysis/Grounds). For example, packaging functions in firms are

done by robots on production lines to increase efficiency and production. Although the use of such new technologies is good for the businesses as it helps them save on costs, increase profits, and work faster, it's bad news for workers since they end up jobless as machines take their jobs (Explanation/Backings). Burt Sharelle (2015), in an article on the impact of technology on people's lives, says that "research indicates that nearly three out of four workers say that advancement in technology is likely to affect their jobs." Therefore, in the end, new technology increases unemployment rates and joblessness hence decreasing the quality of life for some people (Evaluation/Modal Qualifiers). Advancement in technology also reduces performance by employees since some spend more time on smartphones and social networks. Hence, new technology can under-taking the human labor.

Secondly, new technologies have enabled countries and companies to manufacture deadly weapons that have the capacity to cause massive destructions (Interpretation/Claims). Ramey Karehka (2012) argues that modern technology has been a major contributor to endless wars (Analysis/Grounds). As countries compete for superiority, they use new technology to manufacture superior weapons, which is dangerous considering how much damage these weapons can cause (Inference/Warrants). Also, sometimes such weapons get into the hands of criminals who use them for selfish reasons (Explanation/Backings). Hence, new technology can be a major contributor to endless wars (Inference/Warrants).

All in all, (Self-regulation/Rebuttals) new technology has many benefits in the world today. It has simplified things and made it possible to do more in less time and cost. New technologies and innovations come up every day, and most are meant to make the world a better place. However, the new technology is not without its disadvantages (Evaluation/Modal Qualifiers). It leads to loss of jobs, detachment from social life, increased risks of war, and over-dependency on technology. With the current trend in technology, robots, and automated systems will take over the world, and this is a great risk to humans.

According to Bruce and Brooks (1987), advancements in technology are always intended to "make things easier, save time, and increase efficiency". Therefore, new technology is beneficial, but people should try to limit its negative impacts (Self-regulation/Rebuttals).

Indicators of Critical Thinking for Scoring Argumentative Writing

No.	Indicators of Critical Thinking	Score
1.	Interpretation (Claim)	4
2.	Analysis (Ground)	2
3.	Inference (Warrant)	2
4.	Explanation (Backing)	2
5.	Evaluation (Modal Qualifier)	3
6.	Self-regulation (Rebuttal)	3

Appendix 8. Students' Essay with Lowest Score

Male Students

Total Score: 8

LGBT

LGBT stands for lesbian, gay, bisexual, and transgender. Moreover, LGBT is used by everyone who is not heterosexual. Nowadays, not only lesbian, gay, bisexual, and transgender, but LGBT also contains more sexual orientations and multiple gender identities (Interpretation/Claims).

As time passes, LGBT performers in Indonesia are continuously increasing from year to year. In 2012, the Indonesian's LGBT performers are range 1 million, including 348.000 of them spread across East Java (Jawa Pos, 2017). In Bogor, it was recorded that there are 4.164 Indonesian's LGBT performers in 2017 (Dedy, 2021). Furthermore, in 2019, the number of Indonesian's LGBT performers is increased to around 7 million (Hasnah; and Alang, 2019) (Analysis/Grounds). Thus, Indonesia copes the increasing amount of LGBT performers with several applicable laws (Inference/Warrants).

Rita Hendawaty Soebagio (2016) explains, that homosexual abuse is included in the category of criminal code (KUHP) that describes adultery as an act of sexual intercourse committed by a married man or woman with an unmarried woman or man (Explanation/Backings). Moreover, homosexual harassment performer needs to be explained in the criminal code of KUHP (Evaluation/Modal Qualifiers). However, if this activity is carried out by adults who like each other, then the punishment cannot be enforced (Inference/Warrants). This is clearly contradicting the Qur'an and Hadith as basic of Islamic application law (Self-regulation/Rebuttals). Homosexuality is also against Pancasila, UU no 1 of 1974 about Marriage and the field of Marriage. As mentioned before, the number of Indonesian LGBT performers is indeed increasing. Even so, Rita said that LGBT is against morals and the law is strictly prohibiting homosexuality (Evaluation/Modal Qualifiers). **Thus, LGBT in Indonesia should not be tolerated.**

Indicators of Critical Thinking for Scoring Argumentative Writing

No.	Indicators of Critical Thinking	Score
1.	Interpretation (Claim)	1
2.	Analysis (Ground)	1
3.	Inference (Warrant)	2
4.	Explanation (Backing)	1
5.	Evaluation (Modal Qualifier)	2
6.	Self-regulation (Rebuttal)	1

Appendix 9. Validation Letter from The First Validator

SURAT KETERANGAN VALIDATOR

Saya yang bertanda tangan dibawah ini:

Nama : Prof. Dr. Didik Santoso, M.Pd

NIP : 196606161994031006

Jab/Gol : Profesor-Pembina Utama Madya/IV.D

Menerangkan bahwa instrument penilaian critical thinking terhadap tulisan argumentatif mahasiswa tersebut dibawah ini:

Nama : Rinaldy Alidin

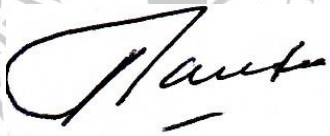
NIM : 202210560211032

Program Studi : Magister Pendidikan Bahasa Inggris

Benar telah selesai divalidasi sesuai dengan indikator yang terdapat pada critical thinking dan indikator penulisan argumentatif, dan indikator pada instrument tersebut telah sesuai dengan penelitian yang berjudul "**Gender Differences in Students' Criticial Thinking in Argumentative Writing"**. Demikian keterangan ini saya sampaikan agar dipergunakan seperlunya demi kepentingan penelitian.

Medan, 13 Mei 2024

Validator,



Prof. Dr. Didik Santoso, M.Pd

NIP. 196606161994031006

Appendix 10. Validation Letter from The Second Validator

SURAT KETERANGAN VALIDATOR

Saya yang bertanda tangan dibawah ini:

Nama : Dr. Sholihatul Hamidah Daulay, S.Ag., M.Hum

NIP : 197506222003122002

Jab/Gol : Lektor Kepala-Pembina/IV.A

Menerangkan bahwa instrument penilaian critical thinking terhadap tulisan argumentatif mahasiswa tersebut dibawah ini:

Nama : Rinaldy Alidin

NIM : 202210560211032

Program Studi : Magister Pendidikan Bahasa Inggris

Benar telah selesai divalidasi sesuai dengan indikator yang terdapat pada critical thinking dan indikator penulisan argumentatif, dan indikator pada instrument tersebut telah sesuai dengan penelitian yang berjudul "**Gender Differences in Students' Criticial Thinking in Argumentative Writing"**. Demikian keterangan ini saya sampaikan agar dipergunakan seperlunya demi kepentingan penelitian.

Medan, 15 Mei 2024

Validator,

Dr. Sholihatul Hamidah Daulay, S.Ag., M.Hum

NIP. 197506222003122002