

## **ANALYZING STUDENTS' UNDERSTANDING OF ARGUMENTATION FALLACIES IN CRITICAL READING INSTRUCTION**

**Agus Setia Budi<sup>1</sup>, Cholimatus Zuhro<sup>2</sup>**

English Study Program, Politeknik Negeri Jember<sup>(1),(2)</sup>

[agus\\_setia@polije.ac.id](mailto:agus_setia@polije.ac.id)

### **Abstract**

This study investigated the Students' understanding of Argumentation Fallacies in Critical Reading of the Students of English Study Program at Politeknik Negeri Jember, especially in the third semester of the 2025/2026 academic year who are taking Critical Reading Class. Using a quantitative descriptive method, the data were gathered through a test that measured students' understanding of errors in giving arguments. The test consists of 55 questions covering eleven types of Argumentation Errors. The test results showed the students had good understanding of Argumentation Fallacies. The most scores were categorized good and very good. In addition, the test had High Internal Consistency, meaning that the test items work well together to measure the same skill and ability. The results of this study show that the students demonstrated a generally strong ability to identify logical fallacies, as reflected in the good mean score and the very high reliability of the test instrument.

**Keywords:** Argumentation Fallacies, Critical Reading Instruction, Students' Understanding

### **Abstrak**

*Penelitian ini menganalisa pemahaman mahasiswa tentang Kesalahan Argumentasi dalam Membaca Kritis pada Mahasiswa Program Studi Bahasa Inggris di Politeknik Negeri Jember, khususnya pada Semester Tiga tahun akademik 2025/2026 yang mengambil Kelas Critical Reading. Dengan menggunakan metode deskriptif kuantitatif, data dikumpulkan melalui tes yang mengukur pemahaman mahasiswa tentang kesalahan dalam memberikan argumen. Tes tersebut terdiri dari 55 pertanyaan yang mencakup sebelas jenis Kesalahan Argumentasi. Hasil tes menunjukkan bahwa mahasiswa memiliki pemahaman yang baik tentang Kesalahan Argumentasi. Skor terbanyak dikategorikan baik dan sangat baik. Selain itu, tes tersebut memiliki Konsistensi Internal yang Tinggi, yang berarti bahwa butir soal tes berfungsi dengan baik untuk mengukur keterampilan dan kemampuan yang sama. Hasil penelitian ini menunjukkan bahwa mahasiswa menunjukkan kemampuan yang secara umum kuat untuk mengidentifikasi kesalahan dalam berlogika, sebagaimana tercermin dalam skor rata-rata yang baik dan reliabilitas instrumen tes yang sangat tinggi.*

**Kata kunci:** Kesalahan Argumentasi, Pemahaman Mahasiswa, Pengajaran Membaca Kritis

### **Introduction**

Critical reading is a higher-order thinking skill that is crucial for students in higher education. Critical reading includes more than just understanding the literal content of a text. It also involves assessing, analyzing, and evaluating the structure and strength of the author's arguments. According to Paul & Elder (2008), critical reading is the art and science of evaluating and analyzing a text. Furthermore, Özensoy (2021) states that critical reading means dividing a text into logical parts, identifying logical

fallacies, and critically evaluating the arguments based on the evidence provided. Generally, students who possess critical reading skills are expected to distinguish between fact and opinion, recognize hidden assumptions, and detect errors in reasoning or argumentation. Therefore, critical reading skills are a key foundation for developing logical and rational thinking in academic settings.

In the context of higher education, the Critical Reading Class plays a strategic role in training students to think analytically and reflectively. Through this course, students are expected to not only understand the content of a text but also be able to evaluate the quality of the arguments presented. As stated by Friedman & Kaganovskiy (2023), in the digital age, critical thinking is crucial because we need to challenge inaccurate or false statements. We must learn to comprehend and assess an argument's core in order to build it. So, one important aspect of critical reading is the ability to recognize errors in argumentation. These errors are often referred to as logical fallacies forms of reasoning that appear convincing but are logically invalid. The ability to recognize logical fallacies helps students avoid being easily influenced by misleading arguments, whether in academic texts, media discourse, or everyday socio-political discussions.

However, based on experience teaching Critical Reading Class, many students still struggle to critically analyze arguments. For example, students often accept seemingly convincing information or claims without evaluating the evidence supporting them, or struggle to distinguish between personal opinions and data-based arguments. Some students also struggle to recognize logical flaws in arguments, such as fallacies in generalization, invalid cause and effect relationships, or misleading analogies. Furthermore, students often struggle to systematically structure their own arguments, making their evaluation of others' arguments less effective. This phenomenon indicates a gap between the learning objectives of critical reading and students' ability to evaluate arguments logically and consistently.

Previous studies on critical reading and argumentative writing in Indonesia have primarily focused on students' literal and inferential comprehension, with relatively few studies examining their ability to identify logical fallacies in texts. For instance, El Khoiri and Widiati (2017) analyzed argumentative essays of EFL students and found that many students committed common logical fallacies, indicating gaps in reasoning skills. Similarly, Selpia (2020) reported that students often struggle to recognize the logical connections within their arguments, leading to weak or unconvincing reasoning. Albiansyah (2019) further highlighted a correlation between students' low critical

thinking disposition and the frequent occurrence of logical errors in their writing. Moreover, Nurhalimah (2014) observed that while students could identify some basic fallacies, they were less able to evaluate persuasive techniques that appeal to emotion. These studies collectively suggest that although critical reading and reasoning skills have been investigated in Indonesia, systematic research specifically measuring students' understanding and recognition of different types of logical fallacies remains limited, highlighting the need for further empirical study.

Related to this, this study was designed to analyze the results of a student comprehension test on eleven types of argumentative fallacies: Non Sequitur, False Cause, Faulty Analogy, Circular Reasoning, Glittering Generalities, Bandwagon, Appeal to Authority, Plain Folks, Straw Man, Personal Attack, and Appeal to Emotion. This test was designed to measure students' ability to recognize logical fallacies based on the examples of arguments provided. Through these test results, researchers can identify the forms of fallacies that are easiest and most difficult for students to recognize. Furthermore, this study will analyze the quality of the test instrument by calculating the difficulty level and discriminatory power of each item, as well as measuring its reliability using the KR-20 coefficient or Cronbach's alpha. This analysis is crucial to ensure that the test used is truly capable of measuring student abilities validly and consistently.

More specifically, the objectives of this study are to measure students' level of understanding various forms of argumentation errors in the context of critical reading, analyze the difficulty level and discriminatory power of each item in the comprehension test, identify the types of thinking errors that are most difficult for students to recognize, and provide pedagogical recommendations to improve the effectiveness of critical reading instruction, particularly in the aspects of logical analysis and argumentation.

This study is highly relevant in the context of developing critical thinking skills in higher education. The results are expected to provide an empirical overview of the extent to which students understand fallacies in argumentation and serve as a reflection for lecturers in designing more effective learning strategies. By identifying the types of errors most frequently overlooked by students, instructors can design learning activities that emphasize argument analysis and logical fallacies detection.

Furthermore, this study is expected to produce a valid and reliable test instrument to measure argumentation-based critical thinking skills. This instrument can be used not only in Critical Reading courses, but also in Academic Writing, Argumentation and Debate, and other courses that emphasize logical reasoning. Thus, this study contributes

to efforts to improve the quality of academic literacy and critical thinking skills among students in higher education.

Conceptually, this study is based on the understanding that critical reading is not only oriented towards understanding the content of the text, but also on evaluating the logic and structure of arguments. Within this framework, errors in giving arguments are seen as an important indicator of logical and reflective thinking skills. Students who are able to recognize errors in reasoning have achieved a higher level of critical understanding, where they not only understand what the author says but also how the author constructs and defends his or her claims. Therefore, mastering the concept of logical fallacies is a key element in developing critical, objective, and rational academic readers and writers.

## **Methods**

This study used a descriptive quantitative approach to analyze students' understanding of various forms of argumentation errors within the context of the Critical Reading Class. This design was chosen because the study focuses on measuring student learning outcomes through systematically developed objective tests and quantitative analysis of item characteristics. This study did not involve treatment or experiments, but rather aims to describe and interpret students' ability to recognize argumentation errors through analysis of test results. Therefore, the results are expected to provide an empirical picture of students' mastery levels and the quality of the instruments used in critical reading instruction.

The subjects of this study were English Study Program students taking the Critical Reading course in the third semester of the 2025/2026 academic year at Politeknik Negeri Jember. The number of respondents in this study was 44 students, consisting of both male and female students with relatively homogeneous academic backgrounds. Participants were selected using a total sampling technique, as all students taking the course served as the research sample. The students who participated in the study had already received material on argument analysis and an introduction to logical fallacies during their studies, and therefore were assumed to have the basic knowledge necessary to answer the comprehension test.

The main instrument used in this study was a test to assess the understanding of logical fallacies. The test was about errors in giving arguments. This test was developed by the researchers based on theoretical studies of logical fallacies and references from various literature sources on critical thinking and argumentation. The test consists of 55

multiple-choice questions, each measuring a specific form of logical fallacy. The eleven types of fallacies measured in this test include: Non Sequitur, False Cause, Faulty Analogy, Circular Reasoning, Glittering Generalities, Bandwagon, Appeal to Authority, Plain Folks, Straw Man, Personal Attack, and Appeal to Emotion. Each type of logical fallacy is represented by five questions, for a total of 55 questions. Each question presents a statement or argument containing a fallacy, and students are asked to choose the type of fallacy that best describes the argument.

The implementation of this study followed a series of structured steps, beginning with the development of a test instrument based on established theories of logical fallacies gathered from various scholarly sources. Each test item was designed to represent a specific type of argumentative fallacy. The instrument was then piloted with a small group of students to assess the clarity of item wording, the appropriateness of the allotted time, and to obtain preliminary validity and reliability estimates. After revisions were completed, the finalized test was administered to 44 students during a 60-minute lecture session under the direct supervision of the researcher. Student responses were subsequently coded in binary form (1 = correct, 0 = incorrect) using a prepared CSV template, and the dataset was analyzed using R statistical software to generate item difficulty, item discrimination, and other relevant measurement indices.

The data in this study were analyzed using descriptive statistics and item analysis. The analysis began by calculating each student's total score based on the number of correct answers, followed by presenting descriptive statistics such as frequency tables, mean, median, mode, and standard deviation. Next, item difficulty was measured using the p-index, which represents the proportion of respondents who answered each item correctly, categorized as easy ( $p > 0.70$ ), moderate ( $0.30 \leq p \leq 0.70$ ), and difficult ( $p < 0.30$ ). Item discrimination was examined using the point-biserial correlation between each item score and the total test score, with criteria classified as good ( $r \geq 0.30$ ), fair ( $0.20 \leq r < 0.30$ ), and weak ( $r < 0.20$ ). Test reliability was calculated using the KR-20 formula for dichotomous data and confirmed with Cronbach's alpha, with reliability considered high if  $\alpha \geq 0.70$ . In addition, the average student performance was calculated for each type of argumentation error to identify the most challenging categories and the results were visualized using bar charts.

*Table 1. The Score Range in Educational Assessment*

Score Range (%)	Interpretation
90 – 100	Very High / Excellent
80 – 89	High / Very Good
70 – 79	Above Average / Good
60 – 69	Average / Moderate
50 – 59	Low
< 50	Very Low

## Result and Discussion

### Result

There were 44 students participated in the *Errors in Giving Argument Test* consisting of 55 items. Table 1 presents the descriptive statistics of the students' overall test performance.

**Table 2.** Descriptive Statistics of Students' Test Scores

Statistic	Value
Number of participants	44
Minimum score	98
Maximum score	38
Mean (M)	77
Median	78
Standard deviation (SD)	13.07
Reliability (KR-20)	0.97

The results indicate that the average score ( $M = 77$ ) is within the good range, suggesting that most students have a good understanding of logical fallacies but still experience difficulties in identifying several types. The reliability coefficient ( $KR-20 = 0.97$ ) shows that the test has very high internal consistency and it is suitable for measuring students' comprehension of reasoning errors.

#### a. Item Test Difficulty and Discrimination

An item analysis was conducted to determine the quality of the test items. Item difficulty is the proportion of students who answered correctly. The test had overall

average  $p = 0.77$ . This shows the test is moderately easy. Related to the distribution of difficulty of the test, it can be found that there were several very easy items ( $p > 0.85$ ) in the test with 41 - 43 correct answers. There were also many moderate items ( $p = 0.30$ – $0.70$ ) in the test and few difficult items ( $p < 0.30$ ) with like Item 20 with 8 correct and 36 incorrect answers. From this findings, it can be concluded that the test has a generally appropriate level of difficulty because it includes a few items that are either too easy or too difficult. Then, related to items discrimination, it require upper–lower group comparison, the wide spread of student scores ( $SD = 13.07$ ) combined with high reliability. It suggests that most items show acceptable to good discrimination and extremely easy or extremely difficult items likely contribute less to discrimination. Furthermore, the findings shows that the test reliability is  $KR-20 = 0.97$ . It indicates very high internal consistency. High internal consistency means that the items in the test work well together to measure the same skill, ability, or construct.

So, based on the result of item analysis, it indicates that: (1) the test is highly reliable and suitable for assessment purposes, (2) most items are at an appropriate difficulty level, (3) a small number of items may need revision due to extreme difficulty levels, and (4) score distribution shows effective differentiation among students.

In addition, it can be found that the test has excellent reliability, balanced difficulty pattern, effective differentiation among students, and strong consistency among items

#### *b. Students' Performance*

The following table presents the students' performance in identifying the types of *errors in giving argument*. It provides information about which types of *errors in giving argument* were most and least understood by the students. The data shows the percentage of the students who answered correctly for each of the eleven fallacy types.

**Table 3.** The Students' Scores by Fallacy Type

Type of Fallacy	Percentage	Interpretation
Non Sequitur	82,3%	High/Very Good
Faulty Analogy	85,6%	High/Very Good
False Cause	70,4%	Above Average/Good
Circular Reasoning	77,1%	Above Average/Good
Glittering Generalities	79%	Above Average/Good
Bandwagon	83%	High/Very Good

Type of Fallacy	Percentage	Interpretation
Appeal to Authority	81%	High/Very Good
Plain Folks	60,2%	Average/Moderate
Straw Man	68%	Average/moderate
Personal Attack	77,1%	Above Average/Good
Appeal to Emotion	86,6%	High/Very Good

From Table 3, it can be seen that students showed the highest performance in identifying Appeal to Emotion (86.6%), Faulty Analogy (85.6%), Bandwagon (83%), Non Sequitur (82.3%), and Appeal to Authority (81%). The students' performance is categorized as "very good". The students performed well in identifying these fallacies because such fallacies are more explicit, direct, and often rely on obvious emotional or logical cues, making them easier to detect.

In contrast to this, Plain Folks (60.2%) and Straw Man (68%) were the most difficult for students. Their performance is categorized as "moderate". The lower performance in identifying Plain Folks due to the subtle nature of this fallacy, which often appears as an ordinary or relatable statement. It makes students hard to recognize the persuasive strategy being used. Similarly, the Straw Man fallacy can be challenging because it requires students to compare the original argument with the distorted version. Many students may fail to notice how the argument is misrepresented, especially when the distortion appears logical or closely related to the initial point.

This finding is supported by studies from Souratia, et al. (2023) and Helwe, et al. (2024), who showed that emotional language is a strong signal for certain types of fallacies. Learners often find it easier to identify fallacies that use emotion or social pressure than those based on more complex logical structures. These fallacies usually include clear language clues, such as emotional words, exaggerated statements, or references to group behavior. Because of these obvious cues, students can quickly connect the statements to familiar persuasive techniques, which helps them recognize the fallacies more accurately.

In contrast, structurally logical fallacies such as Plain Folks and Straw Man require more advanced analytical skills. The Plain Folks fallacy is difficult to spot because the speaker tries to sound like a normal, everyday person to gain trust, and this often seems natural in regular conversation. As a result, students may not notice that it is a persuasive

trick. The Straw Man fallacy is also challenging because students must notice how the original argument has been changed or simplified. To identify it, they need to understand the first argument, compare it with the new version, and judge how it was twisted. These steps take more effort, especially for intermediate-level readers.

Consequently, both earlier research and the current findings show that students depend more on clear emotional clues than on deeper logical thinking when judging arguments. This is why fallacies with obvious signals are easier for them to spot, while more subtle or structurally changed fallacies are harder to recognize.

*c. Distribution of Students' Scores*

**Figure 1.** Correct and Incorrect Answer per Question

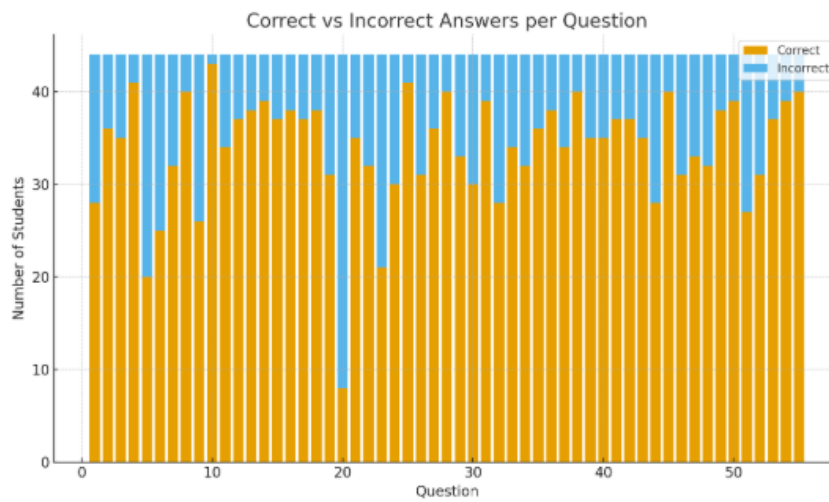


Figure 1 presents the distribution of correct and incorrect responses for all 55 test items. The stacked bars illustrate the number of students who answered each item correctly compared to those who answered incorrectly. Overall, the pattern indicates that most items were successfully answered by a large majority of students.

Across the test, the majority of items show a dominant proportion of correct responses, suggesting that these items were generally easy and well understood by the learners. This indicates that the test content was mostly accessible and aligned with the students' level of proficiency. Items with high correct rates also contribute to the wide spread of total scores, which supports the high reliability obtained through the KR-20 calculation. However, several items exhibited a noticeably lower percentage of correct answers. These items, which appear with a larger portion of incorrect responses, may represent more challenging content or may require revision. Such items might contain

ambiguous wording, higher level reasoning demands, or less familiar concepts for the learners. Their lower performance indicates that students may need additional instruction related to those specific skills or fallacy types.

The overall consistency of the total bar heights indicates that most students attempted all items, minimizing the risk of missing data or irregular response patterns. This uniformity strengthens the validity of the item analysis process.

In general, the chart shows that while the majority of items functioned effectively, a small number require closer examination. These findings support the conclusion that the test demonstrates strong overall quality, with high internal consistency and mostly well-performing items, while also highlighting areas where refinement may enhance future assessments. The frequency distribution indicates that student scores were approximately normally distributed, with most scores clustering around the mean. This normal distribution suggests that the instrument effectively differentiates between high and low performers, confirming its usefulness for diagnostic purposes.

## **Discussion**

The analysis reveals several significant findings:

### *a. Overall Performance*

Students' mean score is 77. It shows that students achieved a good mean score, reflecting a solid and consistent mastery of logical fallacy identification across the test items. This suggests that most students were able to understand, recognize, and interpret the fallacies presented in the assessment, indicating effective instruction and a generally strong level of critical-thinking competency within the group.

Students' better performance in Appeal to Emotion (86.6%), Faulty Analogy (85.6%), Bandwagon (83%), Non Sequitur (82.3%), and Appeal to Authority (81%). It indicates that students are more adept at identifying fallacies that contain explicit emotional, social, or surface-level logical cues, making these types of fallacious reasoning easier to recognize. This pattern supports Browne and Keeley's (2019) argument that learners often rely on emotional cues rather than formal logic in evaluating arguments, suggesting that students tend to notice persuasive elements that are more obvious or affective in nature, while subtler structural reasoning errors require deeper analytical processing.

### *b. Item Functioning and Reliability*

Test items with appropriate difficulty level, together with the very high KR-20 reliability value of 0.97, indicates that the test effectively measures students' ability to

analyze fallacies in a consistent way. The KR-20 value shows that the items work well together and measure the same underlying skill, while the test's difficulty level and item distribution ensure it can clearly differentiate students based on their levels of understanding. In this context, the construct being measured is students' ability to recognize and analyze logical fallacies. Furthermore, consistent means that students' performance does not fluctuate randomly. The test items work together in a predictable way. So, it means that the test consistently measures the same core ability of the students' skill in identifying and analyzing logical fallacies without being influenced by irrelevant factors. This is supported by the high KR-20 value, which indicates strong internal reliability.

*c. Pedagogical Implications*

These findings highlight the need for explicit instruction on logical structures, not just on reading content. It means that students should not only learn the ideas or information in a text, but also be taught how arguments are built, including premises, conclusions, and how ideas connect logically. In other words, teachers must focus on helping students understand how reasoning works, not just what the text says. Activities such as argument mapping, reasoning reconstruction, and peer analysis of fallacious statements can strengthen students' critical reasoning. As suggested by Facione (2020) and Halpern (2014), structured analytical exercises significantly enhance critical thinking when integrated into reading instruction.

Facione (2020) emphasizes that critical thinking improves when learners are guided through structured analytical processes, such as identifying assumptions, evaluating evidence, and distinguishing between strong and weak arguments. He argues that these skills do not develop automatically through reading alone; they must be explicitly modeled and practiced. This supports the idea that students benefit from intentional instruction that makes the logic of arguments visible and understandable. Similarly, Halpern (2014) demonstrates that critical thinking strengthens when students engage in deliberate practice using targeted reasoning tasks. She found that activities requiring learners to analyze argument patterns, detect errors in reasoning, and apply logical principles lead to measurable gains in critical-thinking ability. Halpern also notes that these skills transfer more effectively when learners receive clear explanations and repeated opportunities to apply reasoning strategies in different contexts. These two findings show that teachers need to include clear reasoning exercises that are not only reading activities, So, the students can get better at finding and understanding logical fallacies.

*d. Alignment with Previous Studies*

These results support earlier studies such as Khartite, et al. (2021) and Simonovic, et al. (2023) showing that students still struggle with abstract logical fallacies unless teachers clearly show how logical reasoning works. Therefore, using fallacy-focused reading tasks in Critical Reading classes can help students move from just understanding a text to being able to judge and evaluate the arguments in it.

Khartite, et al. (2021) confirmed that fallacy-focused instruction is an effective pedagogical strategy. Explicit training not only enhances students' ability to detect logical flaws including subtle or structurally complex fallacies—but also fosters deeper metacognitive awareness of reasoning processes. The study underscores the importance of integrating reasoning and fallacy instruction into academic curricula to support students' growth as critical readers, thinkers, and communicators.

In addition, Simonovic, et al. (2023) concluded that structured critical-thinking tasks, especially those that model reasoning processes, play an essential role in strengthening argument analysis skills. Their findings support the growing evidence that deliberate practice, explicit instruction, and guided examples significantly enhance learners' ability to think critically and reason logically. This aligns with the broader argument that critical thinking does not develop merely through exposure to texts; it must be taught and practiced systematically. In conclusion, their studies provide strong evidence that explicit, structured instruction is essential for developing students' critical thinking and argumentation skills. Both studies demonstrate that learners significantly improve when they are guided through reasoning processes, given targeted practice, and taught how to identify and evaluate common fallacies. Importantly, these interventions not only enhance students' ability to recognize flawed arguments but also strengthen their metacognitive awareness and ability to construct logically sound reasoning on their own.

Overall, the findings reinforce the idea that critical thinking is not acquired automatically through reading or general academic exposure. Instead, it must be intentionally taught through well designed instructional strategies that model logical reasoning, offer guided examples, and provide repeated opportunities for practice. Integrating such structured reasoning activities into academic curricula can lead to meaningful improvements in students' analytical abilities and their capacity to think critically across contexts.

## Conclusion

The results of this study show that the students demonstrated a generally strong ability to identify logical fallacies, as reflected in the good mean score and the very high reliability of the test instrument. The item analysis also confirmed that most test items functioned well, with appropriate difficulty levels and strong discrimination, indicating that the assessment effectively measured students' understanding of reasoning errors. The performance across fallacy types further revealed that students were more successful in recognizing fallacies with clear emotional or surface cues, such as Appeal to Emotion, Faulty Analogy, and Bandwagon. In contrast, fallacies requiring deeper structural analysis, such as Plain Folks and Straw Man, were more challenging. These patterns align with previous studies, showing that learners rely more on obvious linguistic or emotional signals than on complex logical reasoning unless explicit instruction is provided. The findings also support earlier research indicating that critical-thinking skills develop more effectively when teachers model reasoning processes, provide structured practice, and offer guided analysis of arguments. Overall, this study confirms that students' ability to recognize fallacies does not develop automatically; it requires intentional, systematic instruction. Therefore, integrating explicit fallacy-focused activities, argument reconstruction tasks, ongoing formative assessments, and interdisciplinary approaches is essential for strengthening students' critical reading and reasoning skills.

## References

- Albiansyah, R. S. D. (2019). Logical fallacies, critical thinking dispositions, and students' argumentation skills (Tesis, UIN Jakarta). <https://repository.uinjkt.ac.id/dspace/handle/123456789/61274>
- El Khoiri, N., & Widiati, U. (2017). Logical fallacies in EFL learners' argumentative writings. *Dinamika Ilmu*, 17(1), 71–81. <https://doi.org/10.21093/di.v17i1.638>
- Facione, P. A. (2020). *Critical thinking: What it is and why it counts* (7th ed.). Insight Assessment.
- Friedman, H. H., & Kaganovskiy, L. (2023). *Logical Fallacies: How They Undermine Critical Thinking and How to Avoid Them*. Research Gate.
- Halpern, D. F. (2014). *Thought and knowledge: An introduction to critical thinking* (5th ed.). Psychology Press.
- Helwe, C., Witon, W., Habernal, I., Ghanem, B., & Nakov, P. (2024). MAFALDA: A benchmark and comprehensive study of fallacy classification. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*. ACL Anthology.
- Khartite, B., & Hellalet, N. (2021). The Impact of Teaching Reasoning Fallacies on the Critical Thinking Ability of Moroccan Engineering Students: The Case of ENSAM

Meknés. *International Journal of Linguistics, Literature and Translation*, 4(9), 222–232. <https://doi.org/10.32996/ijllt.2021.4.9.23> [researchgate.net+2al-kindipublishers.org+2](https://www.researchgate.net/publication/351111111)

- Nurhalimah, N. (2014). Critical thinking skills, informal fallacy, and argumentative writing in Indonesian students (Disertasi, UIN Jakarta). <https://repository.uinjkt.ac.id/dspace/handle/123456789/58762>
- Özensoy, A. U. (2021). Critical reading as analysing written and visual materials according to objective criteria, distinguishing facts and opinions ... *Eurasian Journal of Educational Research*, 93, 319–336.
- Paul, Richard, and Linda Elder. 2008. *The Miniature Guide to Critical Thinking: Concepts and Tools*. Dillon Beach, CA: Foundation for Critical Thinking.
- Selpia, D. (2020). Students' reasoning and fallacies in Indonesian EFL argumentative writing. Atlantis Press. <https://www.atlantis-press.com/article/125938667.pdf>
- Simonović, B., Vione, K. C., Stupple, E., & Doherty, A. (2023). It is not what you think it is how you think: A critical thinking intervention enhances argumentation, analytic thinking and metacognitive sensitivity. *Thinking Skills and Creativity*, 49, Article 101362. <https://doi.org/10.1016/j.tsc.2023.101362>
- Souratia, Z., Aengus, O., Mosbach, M., Houlsby, N., & Ruder, S. (2023). Robust and explainable identification of logical fallacies in natural language arguments. Swiss Federal Institute of Sport Magglingen ([cydcampus.admin.ch](http://cydcampus.admin.ch)).
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55.