PLAGIARISM IN THE INDONESIAN HIGHER EDUCATION: A STUDY OF NON-ENGLISH MAJOR UNDERGRADUATES

By Dang Hartono

PLAGIARISM IN THE INDONESIAN HIGHER EDUCATION: A STUDY OF NON-ENGLISH MAJOR UNDERGRADUATES

Dang Arif Hartono*[1] & Stefanus Angga Badara Prima[2]
[1]dang.arif@podomorouniversity.ac.id, [2]stefanus.angga@podomorouniversity.ac.id

Universitas Agung Podomoro

ABSTRACT

This study aims to investigate the prevalence of plagiarism among university students and identify the specific types of plagiarism that are most commonly committed in their written work. Thirty-nine essays written by third-year non-English-major university students were analyzed using iThenticate software and were subsequently manually analyzed to identify the frequency and types of plagiarism present in the essays. The results of the analysis reveal that the students' essays contained a relatively low level of plagiarized content, comprising only 9.87% of the total text. The most prevalent form of plagiarism detected was mosaic plagiarism, with 39 cases identified. This was followed by verbatim plagiarism, with 16 cases identified. The findings of this research suggest that the students may have difficulty dealing with mosaic plagiarism and may require comprehensive training on proper paraphrasing techniques.

Keywords: plagiarism, plagiarism checker, iThenticate

A. INTRODUCTION

Higher education institutions require students to demonstrate their understanding of course material through various assignments, such as essays, reports, and papers. These assignments typically become more challenging as students progress through their studies. In Indonesia, for example, students must write a thesis to graduate, which applies to both undergraduate and postgraduate students. To prepare students for this, universities and colleges offer academic writing and research methodology courses as part of their curriculum. The focus of these courses and how they are delivered can vary depending on various factors. However, at least one or two sessions are typically devoted to discussing academic misconduct, including plagiarism.

Despite efforts to prevent it, plagiarism remains prevalent in Indonesian colleges and universities. Patak et al. (2021) believe it is a widespread practice in these institutions. Similarly, Akbar and Picard (2019) view plagiarism as a "critical issue that hinders development and innovation" in the Indonesian academic community.

Research has provided some insight into understanding the causes of plagiarism. According to Jereb et al. (2018) the high incidence of plagiarism among university students is primarily due to greater access to information and technology. Other researchers, such as Bloch (2012), Ehrich et al. (2016), and Kayaoğlu et al. (2016), suggest that cultural norms and values may also play a role. Ehrich et al. (2016) further argue that students from some Asian countries may not view plagiarism as morally questionable, as rote memorization is often emphasized in education systems in those countries. Additionally, Doss et al. (2016) point out that insufficient linguistic competence and inability to articulate ideas or concepts can also lead to plagiarism among students.

While previous studies have helped understand the causes of plagiarism, research on the severity and types of plagiarism committed by university students is limited, particularly in the Indonesian context. One study addressing this topic in the Indonesian context was conducted by Sulaiman and Sulastri (2018). Utilizing essays written by 44 English-major students, they found that verbatim plagiarism was the most common type. However, this study has a significant limitation. The study utilized a free version of an online plagiarism checker (Duplichecker) with a 1,000-word limit for each check, which limits its ability to detect plagiarism. Additionally, it is unclear whether Duplichecker compares documents against an extensive database of scholarly articles and theses, raising questions about its rigor as a plagiarism checker service.

This study is set to investigate and uncover the most prevalent forms of plagiarism among essays written by undergraduate students and assess the degree of their severity, thus addressing a gap in the literature. Specifically, this study focuses on non-English-major students, as there is little research on this group in Indonesia.

B. METHOD

1. Context and Participants

The present study was conducted at a private university in Jakarta, Indonesia. Many lecturers at this institution hold graduate degrees from overseas universities in Australia, the United States, and Taiwan. Several of these lecturers have expressed concerns about instances of plagiarism that they have encountered within their classes.

In total, 39 students participated in this study. Of these, 22 were female, and 17 were male. They were drawn from two different academic programs: nine students were majoring in law, while thirty were majoring in product design. Most of these students were in the fifth semester of their studies, as their curricula required them to take an academic writing course during their third year.

The students participated in a compulsory academic writing course. One session of this course was dedicated to discussing various types of plagiarism and how to avoid it through proper paraphrasing and citation. For half a semester, the students focused on writing an argumentative essay and were required to submit their essays for mid-semester evaluation.

2. Data Analysis

The essays were subsequently analyzed for plagiarism using iThenticate, a professional, subscription-based plagiarism checker specifically designed for researchers, scholars,

publishers, and universities. iThenticate compares documents against scholarly databases (e.g., ProQuest, Elsevier, etc.), theses, and open content found on the internet to detect instances of plagiarism. The software generates a report indicating the similarity percentage that a document shares with various sources on the internet. The higher the percentage, the greater the likelihood that the document contains plagiarized content. However, to ensure accuracy, the reports generated by iThenticate were manually reviewed to avoid false alarms or inaccurate detections.

Quantitative data from the reports were analyzed statistically to measure the severity of plagiarism cases. Confirmed cases of plagiarism were also categorized into types of plagiarism set by (Harvard College, 2023), as previous studies (e.g., Gullifer & Tyson, 2014; Gottardello et al., 2017) have found that plagiarism can be defined differently and can be confusing. This classification was chosen as it eliminates room for multiple interpretations.

The classification of plagiarism proposed by Harvard College (2023) includes six types: verbatim plagiarism, mosaic plagiarism, inadequate paraphrasing, uncited paraphrasing, uncited quotation, and using material from another student's work. Verbatim plagiarism, also known as copy-and-paste plagiarism, is the act of reproducing the language of an original source, word for word, without proper acknowledgement through citation (Harvard College, 2023). This form of plagiarism is a serious academic offense, as it involves the direct duplication of source material without giving credit to the original author.

Mosaic plagiarism refers to the act of using words, phrases, or sentences from multiple sources without proper citation (Harvard College, 2023). This form of plagiarism occurs when an individual combines various sources, paraphrases them, and presents the work as their own.

Inadequate paraphrasing is the act of using language from an original source, but changing only a few words or phrases without fundamentally altering the structure of the text (Harvard College, 2023). This form of plagiarism is a serious academic offense as it involves the misrepresentation of authorship, failure to give credit to the original source and misleading the audience into thinking the paraphrase is original work.

Uncited paraphrasing and quotation refer to the use of language from an original source without proper acknowledgement through citation (Harvard College, 2023). While the paraphrasing or quotation may be well done, the failure to give credit to the original source constitutes academic misconduct. These actions are often viewed as less severe because they may be a result of a lack of knowledge about academic writing conventions.

C. FINDINGS AND DISCUSSION

1. Findings
The average word count for the essays was 794.44, with a standard deviation of 248.88. The maximum word count was 1751, while the minimum was 205. These results suggest that the students were given a sufficient amount of time and opportunity to produce original work. Descriptive statistics for these findings can be found in Table 1.

Table 1. Descriptive statistics

	N	Min	Max	M	SD
Word count	39	205	1,751	794.44	248.88
Percentage of plagiarized content	39	0	59	9.87	12.02

The analysis of iThenticate reports revealed instances of plagiarism in varying degrees. The average percentage of plagiarism identified was 9.87%, with a standard deviation of 12.02%. Overall, these findings suggest that the submitted essays contained a relatively small number of plagiarized contents. However, it should be noted that a number of essays were found to contain confirmed instances of severe plagiarism.

In addition to analyzing the essays as a whole, the researchers also looked into any potential differences between the male and female students. The results showed that the female participants had an average word count of 851.55, with a standard deviation of 219.40, while the male participants had an average word count of 720.53, with a standard deviation of 271.38. In terms of plagiarism, the female participants had an average percentage of 9.36%, with a standard deviation of 10.04%, and the male students had an average percentage of 10.53%, with a standard deviation of 14.50. These results suggest that while female students tended to produce longer essays, there was no significant difference in the percentage of plagiarism between male and female participants.

Our analysis also revealed that the most prevalent form of plagiarism detected was mosaic plagiarism, with 39 cases identified. This was followed by verbatim plagiarism, with 16 cases identified. Additionally, we identified 7 cases of inadequate paraphrase. However, it should be noted that no instances of uncited paraphrasing, uncited quotation, or copying another student's work were identified in our sample.

Another important consideration that emerged from our comparison and categorization of iThenticate reports is that the percentages displayed on the reports should not be regarded as definitive. For instance, in one essay, iThenticate reported a similarity index of 13%, which is a relatively high percentage for an essay of 681 words. However, upon close examination, it was discovered that these detections were mostly false positives (e.g., short phrases consisting of fewer than five words, titles of articles, or names of academic journals). These findings underscore the importance of carefully interpreting and evaluating the results generated by plagiarism detection software.

Finally, our analysis of iThenticate reports has led to the identification of a pattern. Specifically, essays with a similarity index of 0-5% have been found to be free of plagiarized materials. However, essays with a similarity index of 6-10% may contain a limited degree of plagiarized materials, which in most instances are considered to be insignificant or false positives. Conversely, essays with a similarity index of 11-20% have been found to contain plagiarized materials of varying degrees of severity. Lastly, essays with a similarity index beyond 20% are likely to contain a substantial proportion of plagiarized materials, whether in the form of verbatim plagiarism or mosaic plagiarism. The data from which this analysis is drawn can be found in Appendix A.

2. Discussion

Our research examines two key issues: the prevalence of plagiarism among university students, and the specific types of plagiarism most commonly committed in their writing. Our findings indicate that, overall, the students' essays contained a relatively low level of

plagiarized content, comprising just 9.87% of the texts they produced. This is a positive outcome. We believe that this relatively low percentage may be attributed to a three-hour session we conducted with the students, which provided an in-depth discussion of plagiarism, paraphrasing, and quoting. This session was held during the fourth week of the semester, approximately one month prior to the deadline for submitting the essays. In other words, the issue of plagiarism was brought to the students' attention well before they began writing, which may have prompted them to make a conscious effort to avoid plagiarism.

Despite the low overall prevalence of plagiarism in the submitted essays, we did identify a number of instances of severe plagiarism, particularly in the form of mosaic plagiarism. This type of plagiarism, also known as patchwork plagiarism, can often be the result of a lack of understanding of how to properly paraphrase and integrate sources into one's own writing. The results of our study indicate that some students may still have difficulty understanding the idea of plagiarism and how to prevent it. This is supported by previous research by Mustafa (2019) and Gullifer and Tyson (2014), who have found that students lack sufficient knowledge or understanding of what constitutes plagiarism.

Our study did not reveal any substantial disparities in the rates of plagiarism among male and female students, which is at odds with the results of the study conducted by Jereb et al. (2018), who found that female students had a less tolerant attitude towards plagiarism in comparison to male students. Our study, however, analyzed students' essays, rather than just measuring their attitudes towards plagiarism. This suggests that while individuals may have intentions to avoid plagiarism, their actions ultimately reveal the reality of their adherence to this ideal.

In our analysis of submitted essays, we also identified sixteen instances of verbatim plagiarism, which varied in both length and severity. This discovery was particularly concerning, given that the students had previously participated in a dedicated 3-hour session on the topic of plagiarism and had been explicitly warned of the potential consequences. Further examination revealed that in these sixteen cases, no attempt had been made to paraphrase the original sources. This suggests that the students likely engaged in intentional plagiarism.

Another significant finding of our study is that plagiarism detection software should be utilized with prudence. Our examination revealed that the percentages reported by the software should not be considered as conclusive, as they may include instances of false positives. This aligns with prior research (Manley, 2021; Stapleton, 2012; Warn, 2006), which also emphasizes the need for caution when interpreting the results of plagiarism detection software. Manley (2021) stresses that although the quantitative data obtained from the software can be informative, a qualitative evaluation is also necessary, where academic staff evaluate whether the plagiarism found was intentional or not.

Finally, our analysis has led to the identification of a pattern in terms of similarity index and plagiarized content. In general, essays with a similarity index of 0-10% were found to be free of plagiarized materials or contained a limited and negligible degree of plagiarism. Conversely, essays with a similarity index of 11-20% were found to contain plagiarized materials to varying degrees of severity. Additionally, essays with a similarity index beyond 20% were found to likely contain a large proportion of plagiarized content. These findings call into question the current standards used by some colleges and universities in Indonesia, which often set a maximum similarity index of 20-25%. Based on our analysis, we suggest

that a similarity index of 10-15% would be a more appropriate standard. However, it is important to note that manual evaluation by faculty members is still necessary to fully determine whether plagiarism is intentional or not.

D. CONCLUSION

In conclusion, our research has highlighted several key issues related to plagiarism among university students in Indonesia. Overall, we found that the prevalence of plagiarism in the essays submitted by students was relatively low. However, we also identified instances of severe plagiarism, particularly in the form of mosaic plagiarism. This implies that some students may still have difficulty grasping the idea of plagiarism and how to avoid it.

One important takeaway from our research is the need for continuous training or special workshops to train students about plagiarism and how to avoid it. Our findings indicate that a 3-hour session is not sufficient to fully equip students with the knowledge and skills needed to avoid plagiarism. Furthermore, our analysis of the similarity index and plagiarized content revealed a pattern that suggests current standards used by some colleges and universities in Indonesia may be too relaxed. We recommend that a similarity index of 10-15% would be a more appropriate standard, but it is important to note that manual evaluation by faculty members is still necessary to fully determine whether plagiarism is intentional or not.

This research has several limitations that should be considered. One issue is that the sample size used in the study is relatively small, and future research should consider increasing the sample size to improve the ability to generalize the findings. Another limitation is that the data used for analysis was only based on the analysis of students' essays and the iThenticate reports, which only provided a limited understanding of plagiarism committed by students. To gain a more complete understanding of students' attitudes and understanding of plagiarism, future studies should also consider using additional methods such as interviews or surveys.

REFERENCES

Akbar, A., & Picard, M. (2019). Understanding plagiarism in Indonesia from the lens of plagiarism policy: Lessons for universities. *International Journal for Educational Integrity*, *15*(1). https://doi.org/10.1007/s40979-019-0044-2

Bloch, J. (2012). Plagiarism across Cultures: Is There a Difference? In C. Eisner & M. Vicinus (Eds.), *Originality, Imitation, and Plagiarism* (pp. 219–230). University of Michigan Press.

Doss, D. A., Henley, R., Gokaraju, B., Mcelreath, D., Lackey, H., Hong, Q., & Miller, L. (2016). Assessing Domestic vs. International Student Perceptions and Attitudes of Plagiarism. *Journal of International Students*, *6*(2), 542–565. http://jistudents.org/

Ehrich, J., Howard, S. J., Mu, C., & Bokosmaty, S. (2016). A comparison of Chinese and Australian university students' attitudes towards plagiarism. *Studies in Higher Education*, *41*(2), 231–246. https://doi.org/10.1080/03075079.2014.927850

Gottardello, D., del Mar Pamies, M., & Valverde, M. (2017). Professors' perceptions of university students' plagiarism : A literature review. *BiD: Textos Universitaris de Biblioteconomia i Documentacio*, *39*. https://doi.org/10.1344/BiD2017.39.12

Gullifer, J. M., & Tyson, G. A. (2014). Who has read the policy on plagiarism? Unpacking students' understanding of plagiarism. *Studies in Higher Education*, *39*(7), 1202–1218. https://doi.org/10.1080/03075079.2013.777412

Harvard College. (2023). *What Constitutes Plagiarism?* Harvard College Writing Program. https://usingsources.fas.harvard.edu/what-constitutes-plagiarism

Jereb, E., Perc, M., Lämmlein, B., Jerebic, J., Urh, M., Podbregar, I., & Šprajc, P. (2018). Factors influencing plagiarism in higher education: A comparison of German and Slovene students. *PLOS ONE*, *13*(8), e0202252. https://doi.org/10.1371/journal.pone.0202252

Jereb, E., Urh, M., Jerebic, J., & Šprajc, P. (2018). Gender differences and the awareness of plagiarism in higher education. *Social Psychology of Education*, *21*(2), 409–426. https://doi.org/10.1007/s11218-017-9421-y

Kayaoğlu, M. N., Erbay, Ş., Flitner, C., & Saltaş, D. (2016). Examining students' perceptions of plagiarism: A cross-cultural study at tertiary level. *Journal of Further and Higher Education*, *40*(5), 682–705. https://doi.org/10.1080/0309877X.2015.1014320

Manley, S. (2021). The use of text-matching software's similarity scores. *Accountability in Research*, 1–27. https://doi.org/10.1080/08989621.2021.1986018

Mustafa, F. (2019). "I think it is not plagiarism": How little do Indonesian undergraduate EFL students understand plagiarism? *Asian EFL Journal*, *21*(2), 74–91. https://www.researchgate.net/publication/331547836

Patak, A. A., Wirawan, H., Abduh, A., Hidayat, R., Iskandar, I., & Dirawan, G. D. (2021). Teaching English as a Foreign Language in Indonesia: University Lecturers' Views on Plagiarism. *Journal of Academic Ethics*, *19*(4), 571–587. https://doi.org/10.1007/s10805-020-09385-y

Stapleton, P. (2012). Gauging the effectiveness of anti-plagiarism software: An empirical study of second language graduate writers. *Journal of English for Academic Purposes*, *11*(2), 125–133. https://doi.org/10.1016/j.jeap.2011.10.003

Sulaiman, R., & Sulastri. (2018). TYPES AND FACTORS CAUSING PLAGIARISM IN PAPERS OF ENGLISH EDUCATION STUDENTS. *Inspiring: English Education Journal*, *1*(1), 95–104.

Warn, J. (2006). Plagiarism software: no magic bullet! *Higher Education Research & Development*, *25*(2), 195–208. https://doi.org/10.1080/07294360600610438

Last name Author-1, Last name Author-2 & Last name Author-3: Title of manuscript is short …

Appendix A- The Percentage and Number of Plagiarism Cases Found

No.	ID	Percentage of plagiarism	Number of Cases Found					
			Verbatim	Mosaic	Inadequate paraphrase	Uncited paraphrase	Uncited quotation	Using material from another student's work
1	Essay 01	18	2	3				
2	Essay 02	2						
3	Essay 03	13						
4	Essay 04	20			1			
5	Essay 05	0						
6	Essay 06	0						
7	Essay 07	16	1					
8	Essay 08	3						
9	Essay 09	29	2	3				
10	Essay 10	2						
11	Essay 11	32	3	2	2			
12	Essay 12	0						
13	Essay 13	7		1				
14	Essay 14	0						
15	Essay 15	0						
16	Essay 16	14	1	1				
17	Essay 17	21		6				
18	Essay 18	12	1					
19	Essay 19	7		2				
20	Essay 20	18		2				
21	Essay 21	14		2				
22	Essay 22	9						
23	Essay 23	1						
24	Essay 24	2						
25	Essay 25	3						
26	Essay 26	5						
27	Essay 27	0						
28	Essay 28	0						
29	Essay 29	0						
30	Essay 30	16		5				
31	Essay 31	59	5	2				
32	Essay 32	8		2				
33	Essay 33	27	1	7				
34	Essay 34	0						
35	Essay 35	6			1			
36	Essay 36	4						
37	Essay 37	1						
38	Essay 38	4						
39	Essay 39	12		1	3			

PLAGIARISM IN THE INDONESIAN HIGHER EDUCATION: A STUDY OF NON-ENGLISH MAJOR UNDERGRADUATES

8%

SIMILARITY INDEX

PRIMARY SOURCES

1	**www.scribd.com** Internet	39 words — 1%
2	**e-journal.stkipsiliwangi.ac.id** Internet	33 words — 1%
3	**www.scilit.net** Internet	18 words — 1%
4	**journalscmu.sinaweb.net** Internet	16 words — 1%
5	**vdocuments.net** Internet	16 words — 1%
6	Minh Ngoc Tran, Linda Hogg, Stephen Marshall. "Understanding postgraduate students' perceptions of plagiarism: a case study of Vietnamese and local students in New Zealand", International Journal for Educational Integrity, 2022 Crossref	12 words — < 1%
7	Williams, Brian A.. "A Mixed Methods Study: Understanding Risk Factors for Plagiarism Among Mba Students at a Northeastern U.S. Research University.", Drexel University, 2021 ProQuest	11 words — < 1%

8	rd.springer.com Internet	10 words — < 1%
9	sajs.co.za Internet	10 words — < 1%
10	Mohamed G. Hussein. "The awareness of plagiarism among postgraduate students at Taif University and its relationship to certain variables", Cogent Social Sciences, 2022 Crossref	9 words — < 1%
11	researchbank.rmit.edu.au Internet	9 words — < 1%
12	s3-ap-southeast-2.amazonaws.com Internet	9 words — < 1%
13	srhe.tandfonline.com Internet	9 words — < 1%
14	Conrad, Nina. "How Students Seek and Use Writing Support: Exploring the Spectrum of Literacy Brokering Practices in Higher Education", The University of Arizona, 2022 ProQuest	8 words — < 1%
15	Mary Ann Alua, Nasir Koranteng Asiedu, Deborah Mwintierong Bumbie-Chi. "Students' Perception on Plagiarism and Usage of Turnitin Anti-Plagiarism Software: The Role of the Library", Journal of Library Administration, 2022 Crossref	8 words — < 1%
16	philpapers.org Internet	8 words — < 1%
17	link.springer.com	