Evaluating the effectiveness of ChatGPT to improve English students' writing skills

Efendi Hidayatullah

Magister Pendidikan Bahasa Inggris, Universitas Sarjanawiyata Tamansiswa.
Jl. Kusumanegara No.157, Yogyakarta 55165, Indonesia
Corresponding Author. Email: effendihidayatullah@gmail.com

Abstract: Artificial intelligence is one of the issues that people are concerned about in daily life, especially in education. ChatGPT (Generative Pre-Training Transformer), also called artificial intelligence, was developed by OpenAI. It is an artificial intelligence that many consider to disrupt the integrity of education. This research aims to describe how students use ChatGPT in English writing skills and avoid plagiarism. The method used in this research is qualitative, and the data collection techniques are observation, interviews, and analysis of test results. The expected result of this study is that students are reasonably integrated into the use of ChatGPT and avoid plagiarism.
Keywords: Artificial intelligence, ChatGPT, Integrity, plagiarism.

INTRODUCTION

Artificial Intelligence is a system with the same intelligence as humans and is characterized by the ability to learn, adapt, solve problems, make decisions, and understand human language. In the development of technology, artificial intelligence (AI) is becoming increasingly popular and widely used in various fields, including education. In education, AI technology can help improve effectiveness and efficiency in the teaching and learning process and provide a more personalized and adaptive learning experience for each individual (Kalla & Smith, 2023). An application of AI technology in education is the use of adaptive learning systems. This system uses data about students' abilities and learning needs to provide learning content that suits students' needs. One technology that has attracted attention is ChatGPT, an artificial intelligence-based language model that can generate text automatically using a data-driven approach (Chen et al., 2023).

Chat GPT (Generative Pre-Training Transformer), also called AI (Artificial Intelligence) developed by Open AI. Chat GPT can assist students in completing tasks or answering questions they may have. Chat GPT is trained to mimic human conversation using a technology called NLP (Natural Language Processing). ChatGPT (Chat Generative Pre-Trained Transformer) uses a conversation format so that humans can easily ask questions to AI-like tools that will automatically get answers quickly, enabling text generation based on user prompts (Halaweh, 2023).

This research presents controversial issues related to the use of ChatGPT in English language education as well as new developments in the use of ChatGPT as a tool in the writing process. In addition, it will also briefly outline the main points to be discussed, including the research methodology to be used, the research results related to the use of ChatGPT in English education, and the problems indicated in the use of ChatGPT by students in English writing skills such as plagiarism and the difficulty in implementing the use of ChatGPT responsibly.

In the Era of Education 4.0, teachers have a great opportunity to provide students with technology to help them with their work, to be trained, not just taught, so that students can find answers to their problems with the help of technology (Savitri, 2019). For tertiary

institutions, adopting a type of technology in student's learning process is a form of an answer to the challenges of lifelong learning needs and the need for over-degree skills.

This research, it is hoped, can contribute to students' understanding of the use of ChatGPT in the context of English language education at Universitas Tidar Magelang. By gaining a deeper understanding of the effectiveness of this technology, effective measures can be found in improving students' writing skills and evaluating the irresponsible use of ChatGPT. Finally, it is hoped that teachers and lecturers are not allergic to the presence of AI technology such as ChatGPT. By trying and using it often, over time you will develop a sense of what the writing patterns produced by ChatGPT are so that more or less you can detect writing made with the help of ChatGPT.

METHODS

This chapter explains the research method, which includes the research design, research subject, the technique of collecting data, place and time of research, research instrument, and data analysis

Research Design

This study uses a descriptive qualitative method to provide an overview of how the student writes text in the GPT Chat system in artificial intelligence.

Research Subject

Researchers chose lecturers and students who took the Paragraph Writing for Academic Context course in the 2nd semester as research subjects to assist researchers in obtaining data.

The technique of collecting data

The data collection techniques used in this study are Observation, interview, and Analysis of Test.

Observations

The researcher uses observation checklists to efficiently focus on classroom English teaching and learning activities.

Interview

The researcher used structured interviews to gather data, and four questions are anticipated to complete the current problem. The following are the questions:

C.2.1. How can students avoid plagiarism while integrating ChatGPT into English writing?

C.2.2. To maintain academic integrity and prevent plagiarism, how do students cite information obtained from ChatGPT?

C.2.3. How can students use ChatGPT as a tool to improve their ability to write English texts without committing plagiarism?

C.2.4. What should students do to balance using ChatGPT as a resource to improve their writing skills and avoid plagiarism?

Analysis of test

Tests to measure students' ability to write English essay tests. In this method, students will be given the task of making an essay of 2 questions by typing and sending via Google Form, which will make it easier for researchers to measure the percentage of the use of GPT chat in the essay test results. The presentation of this data will include data such as the number of

respondents and the results of the test analysis presented with a percentage. This activity will be tested twice by giving two essay tests with a vulnerable time of 2 weeks. The aim is to find out the percentage and comparison of previous test results so that it is known that the use of Chat GPT in English assignments has decreased or increased.

Place and Time of Research

Research Place

The research was conducted at Tidar University Magelang. This university is a higher education institution located at Jl. Kapten Suparman No.39, Potrobangsan, Kec. Magelang Utara, Magelang City, Central Java, post code 56116.

Research Time

This research will take 2 months, including observations, interviews, and document studies, until reporting the results of the research.

Research Instrument

the instruments used in this research are as follows:

Observation.

The instruments used in observation are checklists, Based on the results of the observation, the researcher uses these rules to check (V) whether or not anything is present, observation sheet, observation guide, and observation guide (observation sheet or observation schedule).

Interview

This instrument includes a description of the research issue provided as a list of questions, similar to how it is presented in point C. 2. The technique of collecting data and interview guide

Analysis of test

A test is a set of questions, exercises, or other tools used to measure the skills, knowledge, intelligence, abilities, or talents of individuals or groups (Arikunto, 2013). This instrument uses exam questions (test or test questions) and Inventory.

In this test, the researcher will provide two simple questions, namely: (1) E.3.1. Write the definition of ESP in 300 words; (2) E.3.2. Write the definition of the thesis in 100 words.

Tools of the test are a laptop or Handphone (optional), wifi or internet, e-mail addresses, and a list of questions. The following is a table of the linkage of Data Collection Methods and Instruments

No	Technique of collet data	Research Instrument
1	Observation	a. Observation sheet
		b. Observation guide
		c. Observation guide (observation sheet or observation schedule)
		d. Checklist
2	Interview	a. Interview guide
		b. Questionnaire (cheklist)
3	Analysis of test	a. Exam questions (test or test questions)
		b. Inventory

Data Analysis Technique

In this research activity, data analysis is carried out interactively and takes place continuously, so that the necessary data is obtained perfectly.

| Reduction Data | → | Display Data | → | Verification Data | → | Conclusion Data |

Figure 1. Analysis Data Illustrate from (Sugiyono, 2015)

In this study, the researcher analyzed the data as follows:

Observation

Classify and analyze the implementation of student's behavior doing English assignments, based on the results of observations of the teaching and learning process in the classroom.

Interview

Classifying and analyzing data based on interviews with English teachers and students. Cause the researcher used structured interviews, the researcher asked structured questions to the English teachers. There were four answers that the researcher asked in structured questions, and other answers were obtained from English teachers and students about the implementation of ChatGPT in academic writing and how to prevent massive plagiarism.

Analysis of test results

In this process, the results of working on the two essay questions will be checked to find out the percentage of ChatGPT usage without paraphrasing or with paraphrasing. Checking the plagiarism of the test results will use tools such as:

Openai. detector

https://huggingface.co/spaces/openai/openai-detector

GPTZero X

https://gptzero.me/

Ai Content Detector

https://writer.com/ai-content-detector/
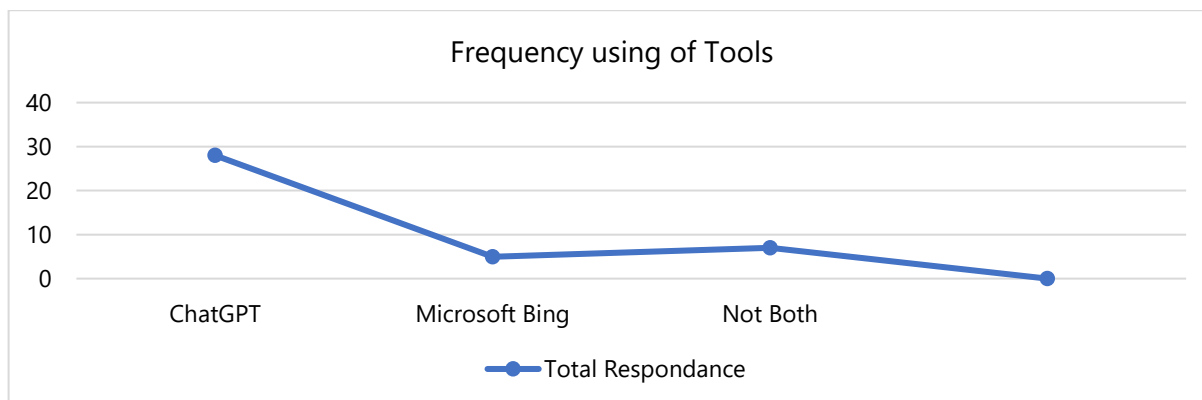
RESULT DAN DISCUSSIONS

Figure 2. The frequency of tools used by students

Figure 1 shows that students who use ChatGPT have a relatively high percentage compared to others, indicating that ChatGPT is still the dominant tool used in text writing. From the results of interviews conducted with 20 respondents

Table 1. Analysis of test

Category test	Question	Respondance
Essay	2	20

Data in the form of sentences was obtained from the test results of 20 students. students were asked to write a reflection in written form. The length of the reflection was one paragraph consisting of about 100 to 150 words. In each reflection, the students produced about 2 answers in the form of typed paragraphs and sent through 2 options, namely by E-mail or Google document. Thus, the total data collected from the test was 80 data.

Table 2. Proposed scoring criteria for essay writing test

Communicative Proficiency	Scoring Weight	Academic Proficiency	Scoring Weight
Intelligibility	5%	Writing in formal academic style including accurate grammar (10%) and vocabulary (10%)	20%)
Comprehensibility	10%	Generating ideas, reasoning, and critiquing	10%
Interpretability	10%	Coherence (and cohesion)	10%
Pertinent content	25%	Organization	10%

Table 3. Scoring Results: Mean Scores of the 3 tool plagiarism detector

Test Taker	Openai. detector	GPTZero X	Ai Content Detector
Student_1	56%	32%	45%
Student_2	26%	32%	11%
Student_3	7%	13%	19%
Student_4	23%	29%	15%
Student_5	29%	22%	17%
Student_6	40%	44%	32%
Student_7	67%	61%	39%
Student_8	34%	29%	23%
Student_9	65%	57%	34%
Student_10	23%	21%	17%
Student_11	9%	15%	13%
Student_12	44%	34%	22%
Student_13	34%	28%	26%
Student_14	22%	29%	22%
Student_15	56%	32%	45%
Student_16	26%	32%	11%
Student_17	7%	13%	19%
Student_18	23%	29%	15%
Student_19	29%	22%	17%
Student_20	40%	44%	32%

Note: Scoring criteria are designed based on the percentage of plagiarism.

This Score is based on the percentage of writing ability, Unlike Smith and Nelson's (1985) description of listening and speaking intelligibility, which is primarily concerned with the prosody of what a person is saying. Intelligibility in writing is the capacity to read the generated text clearly and accurately, this does not refer to the aesthetics of the writer's handwriting.

Comprehensibility refers to the fact that a sentence is understandable or not, and interpretability means if test-takers have been able to make themselves and their intentions understood. We expect that all these aspects of communicative proficiency are easy to evaluate and may not put much burden on the shoulders of raters. If raters can understand test-takers written discourse in the three aspects of understanding, they can easily give the 25 percent scoring points allocated.

As mentioned earlier, each essay is assessed by adopting assessment criteria based on the percentage of plagiarism, communicative proficiency, and academic proficiency. which will be checked for plagiarism using three different tools. The average scores of the three raters can be seen in Table 3.

Based on the data in Table 3, the percentage of plagiarism results is obtained, so the interpretation is based on how much plagiarism is done by using 3 different tools and what is the correlation of the results of checking the 3 tools in question..

CONCLUSION

Figure 1 shows the findings during interviews with students and lecturers about the use of tools to help improve writing skills, namely that students not only use ChatGPT as a tool to help in writing but there are others such as Microsoft Bing, Perflexity, and Sriwijaya Cyber. Another finding is that the perplexity and Sriwijayacyber writing often cannot be detected by plagiarism-checking tools.

The instructional objectives of this study were to develop students' English language skills and motivate students to maintain academic integrity. However, the students were more concerned with the convenience of not having to work hard to find the answers to the prompts on the questions. The students paid less attention to the improvement of their writing skills. The instruction is to not massively use tools or artificial intelligence that are thought to be more helpful than searching for journals, so students do not realize that the main purpose of writing is to develop writing skills, thinking, problem-solving, writing according to writing rules, and being integrated into themselves. It is recommended that teachers inform students about the impact of good and correct writing skills so that it is always part of educational integration.

ACKNOWLEDGMENTS

REFERENCES

Arikunto, S. (2013). *Prosedur penelitian suatu pendekatan praktik*. Rineka Cipta.

Chen, Y., Jensen, S., Albert, L. J., Gupta, S., & Lee, T. (2023). Artificial Intelligence (AI) student assistants in the classroom: Designing chatbots to support student success. *Information Systems Frontiers*, *25*(1), 161–182. https://doi.org/10.1007/s10796-022-10291-4

Halaweh, M. (2023). ChatGPT in education: Strategies for responsible implementation. *Contemporary Educational Technology*, *15*(2).

Kalla, D., & Smith, N. (2023). Study and analysis of Chat GPT and its impact on different fields of study. *International Journal of Innovative Science and Research Technology*, *8*(3).

Savitri, A. (2019). *Revolusi industri 4.0: mengubah tantangan menjadi peluang di era disrupsi 4.0*. Penerbit Genesis.

Sugiyono. (2015). *Metode penelitian pendidikan; pendidikan kuantitatif, kualitati*. Alfabeta CV Bandung.