

# Analysis Of User Experience Of ChatGPT And Gemini Users Using The User Experience Quistionnaire (UEQ) For Education

Ilham Nasrul<sup>[1]\*</sup>, Angraini<sup>[2]</sup>, Muhammad Luthfi Hamzah<sup>[3]</sup>, Eki Saputra<sup>[4]</sup>

Department of Information System Faculty of Science and Technology<sup>[1], [2], [3], [4]</sup>

Universitas Islam Negeri Sultan Syarif Kasim

Pekanbaru, Riau

[12050313471@students.uin-suska.ac.id](mailto:12050313471@students.uin-suska.ac.id)<sup>[1]</sup>, [Angraini@uin-suska.ac.id](mailto:Angraini@uin-suska.ac.id)<sup>[2]</sup>,

[Muhammad.luthfi@uin-suska.ac.id](mailto:Muhammad.luthfi@uin-suska.ac.id)<sup>[3]</sup>, [Eki.saputra@uin-suska.ac.id](mailto:Eki.saputra@uin-suska.ac.id)<sup>[4]</sup>

**Abstract**— AI is becoming more and more crucial in the digital age to support kids in overcoming obstacles to learning and succeeding academically. The use of chatbots is one example of AI progress. Two well-known chatbots are Gemini and ChatGPT. Because they are useful and support a variety of learning tasks, including answering questions, producing articles, expanding knowledge, and other academic activities, both applications are highly well-liked and preferred by students. By using a case study on the Facebook community with the number of samples needed in this study as many as 377 respondents based on the Krejcie and Morgan formula, The purpose of this study was to determine whether user experiences with different applications differed. User experience measurement was carried out using the User Experience Questionnaire (UEQ) approach on the variables of Efficiency, Novelty, Attractiveness, Stimulation, Perspicuity, and Dependability. The results of the study show that all user experience variables for the ChatGPT and Gemini applications received poor ratings, and there were no significant differences in any of these variables. However, based on UEQ measurements, it was found that both applications received better scores on the stimulation and novelty variables, while the attractiveness, clarity, efficiency, and accuracy variables received poor results. To improve user experience in the ChatGPT and Gemini applications, the quality of all variables needs to be enhanced.

**Keywords**— Artificial Intelligence (AI), ChatGPT, Gemini, User Experience, User Experience Questionnaire (UEQ)

## I. INTRODUCTION

In the digital era, technology is increasingly developing and provides convenience for doing various things, including in terms of learning [1]. Artificial intelligence has made a substantial contribution to the advancement of educational technology, enabling more efficient and customized instruction [2]. Technology has a deep impact on our lives and the recent development of artificial intelligence (AI) is part of it. AI is persistently advancing and being applied in various industries [3]. Artificial intelligence (AI) applications (e.g., facial recognition systems, digital assistants, chatbots, recommender systems), proposed in line with the gradual development of technology and the changing needs accordingly, are used in many sectors such as finance, health,

economy, law, medicine, tourism, occupational health and safety, and education [4]. An artificial intelligence (AI) chatbot named ChatGPT uses natural language processing to produce conversational dialogue [5]. AI technologies, exemplified by ChatGPT, have exhibited considerable potential in reshaping the methods through which students acquire knowledge and interact with information [6]. AI systems can be trained to simulate the human brain and carry out routine work using large amounts of data [7].

One form of AI development is the existence of chatbots. A chatbot is an application or service that interacts with users through text conversations. It responds to sentences written by the person at the other end, understanding, learning, and interacting much like a human. This is made possible by Artificial Intelligence or artificial intelligence [8]. The presence of ChatGPT, a chatbot platform developed by OpenAI, continues to be a serious topic of discussion. This is largely due to the side effects of using this artificial intelligence-based platform. With its intelligence capabilities, ChatGPT can assist in facilitating tasks related to text or writing, such as composing letters, copywriting, writing essays, papers, books, poetry, and even academic works like undergraduate theses, master's theses, dissertations, and scientific articles. ChatGPT is one of many generative artificial intelligence (AI) tools, working like a chatbot to generate text responses to user-provided prompts. Like the calculator, the emergence of ChatGPT has quickly created controversy in the educational landscape [9]. It's not wrong to refer to ChatGPT and Large Language Models (LLMs) as a new form of disruption in the academic world. ChatGPT is expected to have a large potential impact on the natural science and engineering domains [10]. An artificial intelligence (AI) research facility in the United States called OpenAI published ChatGPT, a chatbot application, in November 2022. (openai.com, 2022). ChatGPT (ChatGPT Playground) (Chat Generative Pre-Trained Transformer) is an AI-based tool developed by OpenAI, which enables texts generation based on user prompts. It is designed to understand natural language and generate intelligent and relevant responses to user queries [11]. This machine utilizes natural language processing (NLP) technology to respond to human questions in the form of text prompts typed into the application. What has amazed many is

that ChatGPT's responses appear well-structured, with coherent relationships between words or sentences, good accuracy, and the ability to remember previous conversations [12]. The release of ChatGPT (Generative pre-trained transformer) in November 2022 gave many educators and institutions cause for alarm as it provided students with access to software potentially able to assist in the writing process of papers and assignments, provide help with literature reviews, identification of research questions and formatting and language reviews of articles [13].

Google created the generative artificial intelligence chatbot Bard (Gemini). It was improved to PaLM and then eventually transformed into Gemini. Originally, it was based on the large language model (LLM) family LaMDA. Bard was created as a direct reaction to ChatGPT from OpenAI becoming extremely popular very quickly. It was first made available in a small capacity in March 2023 and met with a lackluster welcome; in May, it was made available in additional countries. LaMDA was created and declared in 2021, but caution prevented its public publication [14]. Google rebranded and enhanced its AI platform from Bard to Gemini. The transformation into Gemini introduces a range of advanced features, expanding Google's AI capabilities to include interactions with text, voice, and images. With dedicated apps for Android users and integration into Google apps for Apple users, Gemini aims to be more accessible and user-friendly.

This positions Gemini as a strong competitor to ChatGPT Plus, offering a richer package at the same price point. Google aims to differentiate Gemini by embedding it with distinct personalities. By understanding user intent and responding with personality, Gemini hopes to carve out a unique space in the AI chatbot landscape (suara.com).

OpenAI launched ChatGPT a year ago, and the company and its product immediately became the biggest thing in AI. Now, Google is aiming to shake up this dominance by introducing Gemini. In the competition between Gemini and GPT-4, Hassabis stated that his product excels dominantly enough (cnnindonesia.com).

Researchers conducted observations and surveys to identify issues experienced by some users of ChatGPT and Gemini. The researchers observed through reviews and comments available on the Play Store regarding the ChatGPT and Gemini apps. They also analyzed discussions on the social media platform X (formerly Twitter), where the content of these discussions focused on the various advantages and disadvantages of each application.

According to Himawan Yusuf, Gemini is better than ChatGPT because the information on ChatGPT is limited only up to the year 2022, whereas Gemini can provide the latest data. Additionally, when asked to make a comparison, Gemini automatically creates a summary table of the differences. Joko Joestar believes that ChatGPT is better than Gemini. When using Gemini for coding-related questions, it provided unclear answers, while ChatGPT was able to offer alternative code and explain where the errors were.

Yenni Kartika states that for grammar correction, ChatGPT provides better answers, whereas Gemini gives excessively improvised responses. However, for casual copywriting, Gemini is better than ChatGPT. According to Hidayatur Rahman, after trying various use cases, Gemini is

better for creative tasks like copywriting, report writing, and so on, while ChatGPT is better for analytical tasks such as coding, data analysis, and more.

The role of Artificial Intelligence in student learning has great potential to change the educational paradigm. With proper and responsible use, Artificial Intelligence can enhance the quality and accessibility of education, allowing students to learn in a more personalized and effective manner [2]. Determining the degree of user satisfaction from a user experience standpoint is the goal of this study. As a result, a study is done on ChatGPT and Gemini's user experiences. Using the User Experience Questionnaire (UEQ) approach, the analysis is conducted.

The comprehensiveness of aspects, hedonic quality, namely attractiveness, and pragmatic quality is the advantage of the User Experience Questionnaire (UEQ). Additionally, a Data Analysis Tool in Excel format is provided to facilitate the measurement of the User Experience Questionnaire (UEQ) [15].

This study is unique in that it is the first to directly compare two artificial intelligence platforms—ChatGPT and Google Gemini—a comparison that is still hardly made. This study provides an organized method of assessing the user experiences of both platforms through the use of the User Experience Questionnaire (UEQ), an instrument that has not been thoroughly examined in prior studies. The research's significant emphasis on assessing user experience in the context of interactions with natural language-based AI models is another innovative aspect. This area of study is becoming more and more significant as AI technology is embraced by a wider range of industries. Another distinctive feature of this study is the generation of quantitative data based on UEQ, which can be used to more objectively and quantifiably highlight the advantages and disadvantages of each platform.

## II. LITERATURE REVIEW

### A. Artificial Intelligence

A computer scientist named Professor John Mc Carthy is known as the figure who introduced the concept of artificial intelligence (AI) in 1956. Artificial intelligence comprises a range of methodologies and strategies, such as fuzzy logic, natural language processing, computer vision, expert systems, machine learning, and combinations of these [16] John McCarthy describes artificial intelligence as the science and art of making intelligent machines, especially making intelligent computer programs or applications. A step toward building computers, robots, apps, or programs with human-like intelligence is artificial intelligence.

### B. Chatbot

Chatbot is an artificial intelligence (AI) program and a model of human-computer interaction. Advances in AI technology enable a chatbot to have human-like communication abilities. In processing human language, chatbots are developed using a program called Natural Language Processing (NLP). This program is a branch of artificial intelligence that allows a machine to recognize, learn, and produce human language. In this context, chatbots act as assistants capable of communicating to meet users'

information needs in various contexts[17]

C. ChatGPT

A chatbot program called ChatGPT was released in November 2022 by the OpenAI research lab in the United States (openai.com, 2022). This device uses natural language processing (NLP) technology to answer questions from users that are entered into the program as text prompts.(Suharmawan, 2023). With ChatGPT, people of all ages and backgrounds may converse fluently in a variety of languages without the need for any prior computer or programming knowledge. (Zhou et al., 2023). ChatGPT can be used to facilitate various tasks, from customer service, entertainment, education, finance, healthcare, and more[18]

D. Gemini

On February 7, 2023, Google publicly announced Bard, their AI technology, in an attempt to challenge OpenAI's well-known ChatGPT-3 language model. The CEO of Google revealed this in a publicly accessible blog post. He defined Bard as a conversational AI tool that offers exceptional solutions and can clarify complex circumstances. Right now, Bard from Google comes with a simplified replica of LaMDA. This is so Bard can reach more users and get more feedback because smaller models usually require less processing power.[19] On February 2, 2024, Google changed Bard's name to Gemini formally. As part of its ongoing commitment to giving everyone direct access to its artificial intelligence (AI), Google is making its finest AI models available to all Gemini users in the supported languages and regions as of this week. Google renamed the product Gemini to more accurately represent this dedication. The best direct connection to Google AI is through Gemini. With Gemini, all of the collaboration features you are accustomed to and have grown to appreciate are still available. Additionally, Google has made UI improvements to improve reading, lessen visual distractions, and streamline navigation. (gemini.google.com).

E. User Experience

According to ISO 9241-210 (2010), A product, system, or service's appearance, accessibility, interface performance, and support from prior experiences are all taken into account when measuring user experience, or UX. Before, during, and after using the product, user experience incorporates emotions, trust, preferences, perceptions, bodily and psychological reactions, behavior, and accomplishments. Frank Guo (2012) explains that a good User Experience should fulfill four elements: Usability, Value, Adoptability, and Desirability.(Junita Maulani et al., 2021) Garret, 2011, states that User Experience (UX) is about how a service or product works when someone comes to use it, not about its internal performance.[20]

F. User Experience Questionnaire (UEQ)

According to Santoso et al. (2016), The User Experience Questionnaire (UEQ) is a part of the standard usability test in order to obtain a favorable picture of the user experience from both usability and experiential elements. The goal of UEQ is to enable users to quickly complete an assessment so that a thorough understanding of their perceptions of a product, including their thoughts and actions, may be obtained.[21]

States that UEQ consists of a questionnaire designed to elicit direct user responses regarding their feelings, impressions, and behaviors when using a product. There are 6 measurement scales in UEQ: Efficiency, Novelty, Attractiveness, Stimulation, Perspicuity, and Dependability.[22]

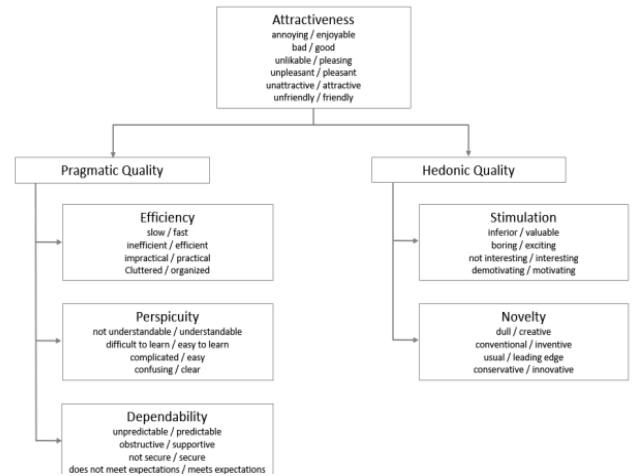


Figure 1. User Experience Questionnaire

UEQ consists of 6 scales with 26 items:

- Attractiveness: overall assessment of the item. Is the product liked or disliked by the user?
- Perspicuity: Is it simple to use and comprehend the product?
- Efficiency: Can users complete their jobs without expending extra energy?
- Dependability: Is the user feeling in charge of the exchange?
- Stimulation: Is using the product exciting and motivating?
- Novelty: Does the product exhibit creativity and innovation? Does the user find it interesting?

G. Previous Research

As for some previous research used as literature review material, such as research.

conducted by [23] In this article, it discusses the intense competition between two leading AIs, Google Bard (Gemini) by Google and ChatGPT by OpenAI. The conclusion of this article is that the major concern for OpenAI and Google Inc. is that other tech giants like Meta are also keeping up with these developments; they are closely monitoring these events and may also be collecting the necessary data to begin studies to develop their own versions of AI-based chat tools. Therefore, Google will undergo challenging tests in the coming weeks or months.

Research from [24] This study will analyze the Gojek and Grab applications with the aim of identifying whether Gojek and Grab have provided a good user experience, as well as which user experience variables from both applications received positive responses from users. Additionally, the study aims to determine if there are any

differences in the experiences felt by users of both applications that lead them to prefer using one application over the other.

Research from [25] Comparing two digital wallet apps with comparable business procedures and user experience features is the goal of this study. The UEQ questionnaire's findings indicate that respondents' perceptions of both digital wallet applications are largely favorable. with the Dana application scoring better in four assessment aspects: Efficiency, Novelty, Attractiveness, Stimulation, Perspicuity, and Dependability. Meanwhile, the OVO application scored better in the aspects of efficiency and stimulation.

Research from [26] In the reviews provided by users for each application, there are various responses from users, including both positive and negative feedback. The purpose of this study is to use the User Experience Questionnaire (UEQ) approach to assess each application's quality. This technique evaluates a product or application's quality by looking at what consumers think of it after they use it. When the UEQ was analyzed, the findings showed that Airy outperformed RedDoorz on a number of UEQ scales.

Research from [27] Politeknik Caltex Riau hosted this study, which included 179 enrolled students as participants. The results of the user experience evaluation using the UEQ showed the following scores for each UEQ aspect: novelty 0.855, stimulation 1.346, perspicuity 1.552, efficiency 1.354, dependability 1.377, and attractiveness 1.375. This suggests that the user experience was well-received in every way. The study's findings can be applied to the construction of a novel student academic information system by incorporating several innovative ideas.

Research from [21] User satisfaction surveys, like the ones intended for students, have not yet been carried out on the LSP UAD website. In order to enhance the information services offered on the website, the purpose of this article is to test customer happiness, or customer Experience (UX). Fifteen students who have used the LSP UAD website as respondents are examined using the User Experience Questionnaire (UEQ) method.

Research from [28] Nearly every industry has been greatly touched by the pandemic. The government is still working to stop the coronavirus from spreading, and one of their efforts is the PeduliLindungi application. This application's goal is to make effort tracking easier. Therefore, to help the government achieve its goal of helping businesses that were halted due to the outbreak to resume, a user experience analysis of the PeduliLindungi application is important. The User Experience Questionnaire (UEQ) tool will be used in the analysis.

Research from [29] TIX ID is a smartphone application that allows users to purchase movie tickets online. It is not necessary for customers to wait in line to buy movie tickets at any time or place. Buying movie tickets is combined with DANA, a third party, as an idea for a digital wallet that is also integrated with several important applications, like Bukalapak and Ramayana. The

purpose of this study is to ascertain how service quality affects users' experiences when using the TIX ID app to buy movie tickets online.

Research from [30] By assessing the scales of beauty, Efficiency, Novelty, Stimulation, Perspicuity, and Dependability, This study will compare the user experiences of the Tiket.com and PegiPegi applications using the User Experience Questionnaire (UEQ) approach. The results show that every category had positive ratings for the Tiket.com application's user experience. The originality rating for the PegiPegi application earned a neutral score, but the scales for attractiveness, perspicuity, efficiency, dependability, and stimulation obtained positive assessments. Tiket.com obtained marginally higher user experience ratings on all UEQ scales when the user experiences of the Tiket.com and PegiPegi applications were compared. None of the scales showed any appreciable variations, though

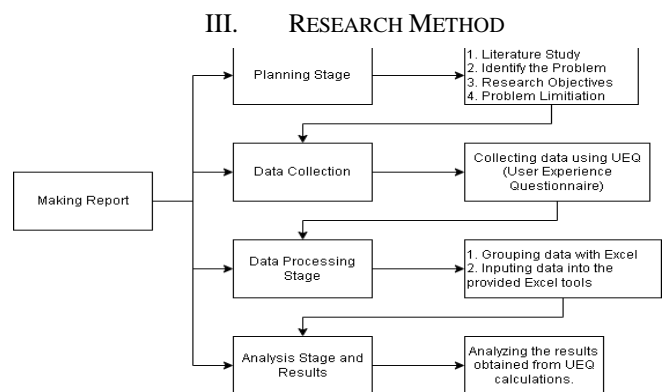


Figure 2. Research Method

The collection of steps that will be taken is known as the research methodology, and this chapter will go into depth about each step. The steps of the research process are shown in the flowchart.

- a) *Planning Stage: Data or information collected directly from research subjects, such as through distributing questionnaires and observations. Furthermore, questionnaires were given to 377 participants to analyze the user experience using questions based on the UEQ questionnaire.*
- b) *Data Processing Stage: Using Microsoft Excel software, descriptive data are reconstructed. Assessing the ChatGPT and Gemini application's user experience with UEQ tools for data processing.*
- c) *Analysis Stage and Results: The goal of the user experience study was to determine how easy it was to use ChatGPT and the Gemini app without encountering any problems or errors, as well as how successful and memorable it was. This will make it easier to recognize and classify the components that need to be developed further or improved upon.*

	1	2	3	4	5	6	7		
annoying	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	enjoyable	1
not understandable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	understandable	2
creative	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	dull	3
easy to learn	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	difficult to learn	4
valuable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	inferior	5
boring	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	exciting	6
not interesting	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	interesting	7
unpredictable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	predictable	8
fast	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	slow	9
inventive	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	conventional	10
obstructive	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	supportive	11
good	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	bad	12
complicated	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	easy	13
unlikable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	pleasing	14
usual	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	leading edge	15
unpleasant	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	pleasant	16
secure	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	not secure	17
motivating	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	demotivating	18
meets expectations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	does not meet expectations	19
inefficient	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	efficient	20
clear	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	confusing	21
impractical	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	practical	22
organized	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	cluttered	23
attractive	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	unattractive	24
friendly	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	unfriendly	25
conservative	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	innovative	26

This is a questionnaire question provided directly by UEQ. This questionnaire is obtained from the UEQ handbook. This questionnaire was distributed to 377 respondents, then the results of the respondents' answers were entered into UEQ Tools in the form of excel for data processing. The sample is part of the population used as the main data in this study. Sampling in this study using Purposive Sampling, this technique selects samples based on certain characteristics relevant to the research.

in this study respondents were taken from a community obtained from Facebook social media. where the community is called "Sharing Tips & Tricks ChatGPT" whose number of members is more than 20 thousand. as for the characteristics of respondents needed in this study are having used ChatGPT and Gemini or one of them, currently a student.

in this study 377 respondents were needed, where this sample size was obtained from the Krejcie and Morgan formula. The reason for using this formula is because it is very useful when researchers have limited resources, be it time, funds, or energy and the Krejcie and Morgan formula is relatively easy to understand and use. Researchers only need to enter several variables such as population and tolerable error rate, then the calculation results will provide an estimate of the number of respondents needed.

Below is the formula of Krejcie and Morgan.

$$n = \frac{X^2 \cdot N \cdot P(1 - P)}{(N - 1) \cdot d^2 + X^2 \cdot P(1 - P)}$$

- $n$  = sample size
- $N$  = Population size
- $X^2$  = chi squared value
- $P$  = population proportion
- $d$  = estimation error

#### IV. RESULTS

After the answers from respondents are processed with UEQ Tools in the form of excel, the results are obtained and explained below. all figures, tables and diagrams are obtained from excel UEQ Tools.

##### A. Analysis of UX value in the ChatGPT Application

Figure 3 shows the results of the questionnaire for each research variable evaluated to calculate the average respondent's answer for each item and the average value.

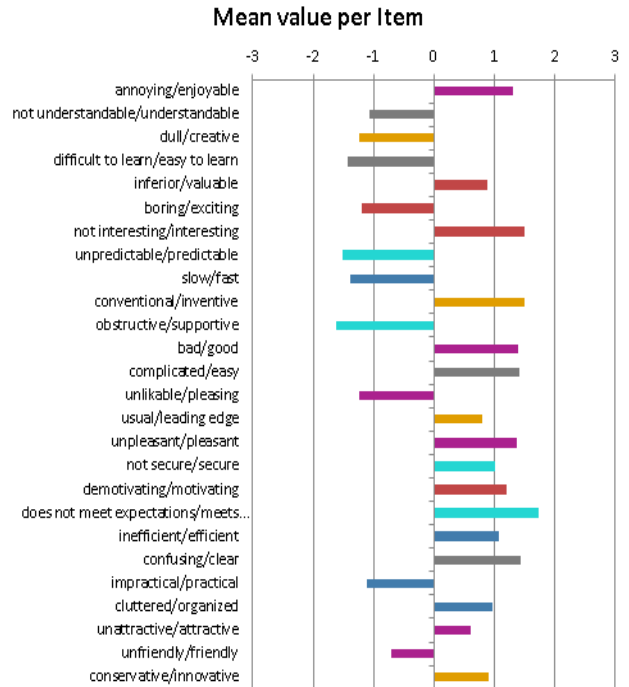


Fig. 1. Mean Value ChatGPT

After obtaining the results for all items from each variable, to consolidate the final results, the mean of each variable is calculated to become the average mean per variable. Additionally, benchmark results are provided to facilitate the visualization of the variable data, as shown in the following figure.

TABLE I. Mean and Variance ChatGPT

UEQ Scales (Mean and Variance)		
Attractiveness	0,454	0,20
Perspicuity	0,088	0,34
Efficiency	-0,118	0,20
Dependability	-0,096	0,37
Stimulation	0,595	0,32
Novelty	0,490	0,46

Table 1 shows the evaluation results in the form of the mean value for each variable measuring user experience using

the UEQ on ChatGPT. These values were obtained from 377 respondents who completed the questionnaire. From these results, it can be seen that the variables Efficiency, Novelty, Attractiveness, Stimulation, Perspicuity, and Dependability are at a neutral evaluation level, with mean values ranging from (-0.2) to (0.6).

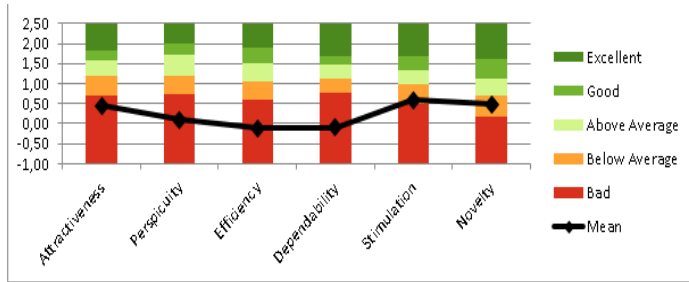


Fig. 2 Benchmark ChatGPT

TABLE II. DATA RESULT CHATGPT

Scale	Mean	Comparisson to benchmark	Interpretation
Attractiveness	0,45	Bad	In the range of the 25% worst results
Perspicuity	0,09	Bad	In the range of the 25% worst results
Efficiency	-0,12	Bad	In the range of the 25% worst results
Dependability	-0,10	Bad	In the range of the 25% worst results
Stimulation	0,60	Below Average	50% of results better, 25% of results worse
Novelty	0,49	Below Average	50% of results better, 25% of results worse

The data in figure 2 and table II depict the benchmark results, indicating that the scales for Attractiveness, Perspicuity, Efficiency, and Dependability received a rating of Bad, meaning they fall within the worst 25% of results. Meanwhile, the scales for Stimulation and Novelty received a Below Average rating, meaning 50% of the benchmark results are better than the evaluated product, while 25% are worse.

B. Analysis of UX value in the Gemini

As shown in Figure 5, the results of the questionnaires administered for each analyzed research variable are as follows: the average of respondents' answers for each item and the average value for each item. Figures and Tables

Mean value per Item

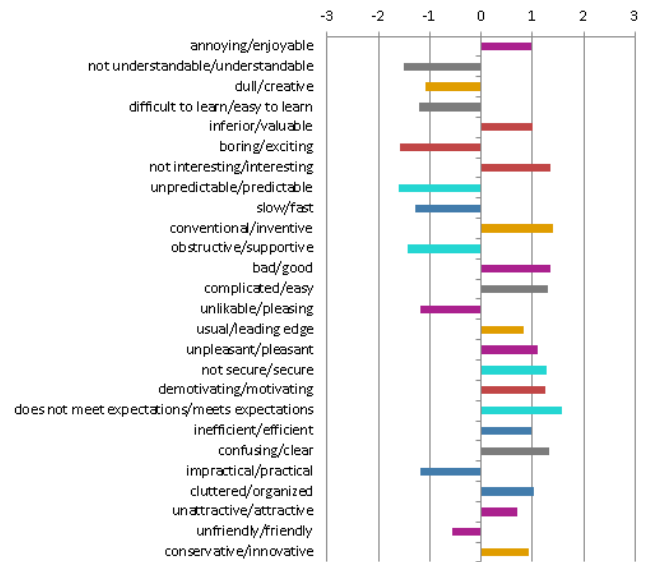


Fig. 3. Mean Value Gemini

After obtaining the results for all items from each variable, to consolidate the final results, the mean of each variable is calculated to become the average mean per variable. Additionally, benchmark results are provided to facilitate the visualization of the variable data, as shown in the following figure.

TABLE III. MEAN AND VARIANCE GEMINI

UEQ Scales (Mean and Variance)		
Attractiveness	0,401	0,22
Perspicuity	-0,021	0,17
Efficiency	-0,115	0,15
Dependability	-0,048	0,20
Stimulation	0,508	0,38
Novelty	0,513	0,43

table 3 shows the evaluation results in terms of the mean value for each variable measuring user experience using UEQ on Gemini. These values were obtained from 377 respondents who completed the questionnaire. From these results, it can be seen that the variables Efficiency, Novelty, Attractiveness, Stimulation, Perspicuity, and Dependability are at a neutral evaluation level, with mean values ranging from -0.1 to 0.5.

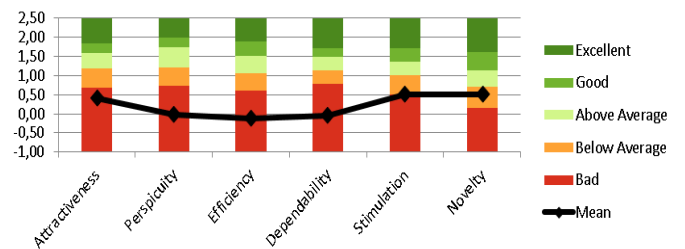


Figure 3. BenchMark Gemini

TABLE IV. DATA RESULT GEMINI

Scale	Mean	Comparisson to benchmark	Interpretation
Attractiveness	0,40	Bad	In the range of the 25% worst results
Perspiciuity	-0,02	Bad	In the range of the 25% worst results
Efficiency	-0,11	Bad	In the range of the 25% worst results
Dependability	-0,05	Bad	In the range of the 25% worst results
Stimulation	0,51	Below Average	50% of results better, 25% of results worse
Novelty	0,51	Below Average	50% of results better, 25% of results worse

The data in Figure 4 and Table IV illustrate the benchmark results, indicating that the scales for Attractiveness, Perspicuity, Efficiency, and Dependability received a rating of Bad, meaning they fall within the worst 25% of results. Meanwhile, the scales for Stimulation and Novelty received a Below Average rating, meaning 50% of the benchmark results are better than the evaluated product, while 25% are worse.

Here are the explanations of each user experience variable for both applications based on the data in figure 2, 4 and table II, IV.

- 1) *Attractiveness of ChatGPT and Gemini applications: The perceived attractiveness value by users of the ChatGPT and Gemini applications is "Bad." This means that the product quality in terms of attractiveness towards users is not good, resulting in users being less interested and not having a positive impression when using the ChatGPT and Gemini applications.*
- 2) *Perspiciuity of ChatGPT and Gemini applications: The perceived perspiciuity value by users of the ChatGPT and Gemini applications is "Bad." This indicates that the clarity and ease of use of the applications are not good and improvements are needed in terms of clarity and user-friendliness.*
- 3) *Efficiency of ChatGPT and Gemini applications: The perceived efficiency value by users of the ChatGPT and Gemini applications is "Bad." This means that the product quality in terms of efficiency is not good, resulting in users being dissatisfied with the efficiency and processing speed of transactions performed by the ChatGPT and Gemini applications.*
- 4) *Dependability of ChatGPT and Gemini applications: The perceived dependability value by users of the ChatGPT and Gemini applications is "Bad." This indicates that the level of reliability of the applications is not sufficiently good and improvement is needed.*
- 5) *Stimulation of ChatGPT and Gemini applications: The perceived stimulation value by users of the ChatGPT and Gemini applications is "Below Average." This means that both applications are capable of providing enjoyment and*

*motivating users to continue using the applications, but improvements are still necessary.*

- 6) *Novelty of ChatGPT and Gemini applications: The perceived novelty value by users of the ChatGPT and Gemini applications is "Below Average." This indicates that both applications have a less innovative and creative appearance, thus not attracting users sufficiently to use the applications.*

The purpose of this study is to assess ChatGPT and Gemini applications' user experiences (UX) using the User Experience Questionnaire (UEQ) technique. The evaluation results show that 4 of the 6 variables assessed get poor results, while the other 2 variables get below average results, meaning that there are still many aspects that must be improved to increase the quality value of each application.

Overall, this research makes an important contribution in understanding user experience on ChatGPT and Gemini applications and identifying areas for improvement.

By using the User Experience Questionnaire method, the researchers were able to capture the users' perspectives more deeply and provide a more in-depth perspective and provide appropriate suggestions for future app improvements.

TABLE V. COMPARISON RESULT

Scale	ChatGPT	Gemini
Attractiveness	0,45	0,40
Perspiciuity	0,09	-0,02
Efficiency	-0,12	-0,11
Dependability	-0,10	-0,48
Stimulation	0,60	0,50
Novelty	0,49	0,51

The table above was created to see a comparison of the results obtained through UEQ analysis, by comparing the value of each aspect obtained from ChatGPT and Gemini. The table shows that ChatGPT excels in four areas: Attractiveness, Perspective, Dependability, and Stimulation. In contrast, Gemini outperforms in Efficiency and Novelty. Based on this comparison, we can conclude that ChatGPT is generally superior to Gemini.

#### CONCLUSION

Based on the research conducted, several conclusions can be drawn: 1) All user experience variables for the ChatGPT application received poor ratings, with "Stimulation" and "Novelty" scoring "Below Average". Additionally, "Attractiveness", "Perspiciuity", "Efficiency", and "Dependability" received "Bad" ratings. 2) All user experience variables for the Gemini application also received poor ratings, with "Stimulation" and "Novelty" scoring "Below Average". Similarly, "Attractiveness", "Perspiciuity", "Efficiency", and "Dependability" received "Bad" ratings.

ChatGPT excels in four areas: Attractiveness, Perspiciuity, Dependability, and Stimulation. Conversely, Gemini outperforms in Efficiency and Novelty. Based on this

comparison, we can conclude that ChatGPT is generally superior to Gemini.

Suggestions for further researchers can combine the UEQ method with other methods such as cooperative evaluation methods, heuristic evaluation, or usability testing. and further researchers can use different evaluation methods so that later comparisons can be made of the results between the methods currently used by researchers and the methods used by subsequent researchers. There are various other AIs that can be used in helping students to improve learning, so further researchers can conduct research with other AI besides Chatgpt and Gemini.

#### REFERENCES

- [1] F. Fauzi, L. Tuhuteru, F. Sampe, A. M. A. Ausat, dan H. R. Hatta, "Analysing the Role of ChatGPT in Improving Student Productivity in Higher Education," *J. Educ.*, vol. 5, no. 4, hal. 14886–14891, 2023, doi: 10.31004/joe.v5i4.2563.
- [2] V. A. Putri, K. Carissa, A. Sotyardani, dan R. A. Rafael, "Peran Artificial Intelligence dalam Proses Pembelajaran Mahasiswa di Universitas Negeri Surabaya," *Pros. Semin. Nas.*, hal. 615–630, 2023.
- [3] S. Pokkakillath dan J. Suleri, "ChatGPT and its impact on education," *Res. Hosp. Manag.*, vol. 13, no. 1, hal. 31–34, 2023, doi: 10.1080/22243534.2023.2239579.
- [4] M. R. Karaman dan İ. Göksu, "Are Lesson Plans Created by ChatGPT More Effective? An Experimental Study," *Int. J. Technol. Educ.*, vol. 7, no. 1, hal. 107–127, 2024, doi: 10.46328/ijte.607.
- [5] R. N. Darma, A. Syahid, Fatma, Ayu Putriana Lestari, Ahmad Basahil, dan Rini Listiya Ningrum, "The Student's Perception of Using ChatGPT for EFL Students," *J. Ilmu Pendidik. Nas.*, vol. 1, no. 3, hal. 143–146, 2023, doi: 10.59435/jipnas.v1i3.193.
- [6] J. Naher, M. H. Maruf, S. T. Bakht, dan S. S. Sadaf, "A Survey to Understand the experience of ChatGPT Usage among Engineering University Students in Bangladesh," *Tech. Rom. J. Appl. Sci. Technol.*, vol. 15, hal. 1–8, 2023, doi: 10.47577/technium.v15i.9660.
- [7] C. K. Lo, "What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature," *Educ. Sci.*, vol. 13, no. 4, 2023, doi: 10.3390/educsci13040410.
- [8] E. L. Amalia dan D. W. Wibowo, "Rancang Bangun Chatbot Untuk Meningkatkan Performa Bisnis," *J. Ilm. Teknol. Inf. Asia*, vol. 13, no. 2, hal. 137, 2019, doi: 10.32815/jitika.v13i2.410.
- [9] A. R. Ellis dan E. Slade, "A New Era of Learning: Considerations for ChatGPT as a Tool to Enhance Statistics and Data Science Education," *J. Stat. Data Sci. Educ.*, vol. 31, no. 2, hal. 128–133, 2023, doi: 10.1080/26939169.2023.2223609.
- [10] L. S. Balhorn, J. M. Weber, S. Buijsman, dan ..., "What does ChatGPT know about natural science and engineering?," *arXiv Prepr. arXiv* ..., hal. 1–18, 2023, [Daring]. Tersedia pada: <https://arxiv.org/abs/2309.10048><https://arxiv.org/pdf/2309.10048>
- [11] C. M. H. Chan *et al.*, "Perception of Harms and Benefits of Electronic Cigarettes Among Adult Malaysian Men: A Comparison by Electronic Cigarette Use and Smoking Status," *Asia-Pacific J. Public Heal.*, vol. 31, no. 7\_suppl, hal. 32S–43S, 2019, doi: 10.1177/1010539519860730.
- [12] W. Suharmawan, "Pemanfaatan Chat GPT Dalam Dunia Pendidikan," *Educ. J. J. Educ. Res. Dev.*, vol. 7, no. 2, hal. 158–166, 2023, doi: 10.31537/ej.v7i2.1248.
- [13] R. G. Jurado, T. Pettersson, dan M. Zwierewicz, "STUDENTS' ATTITUDES TO THE USE OF ARTIFICIAL INTELLIGENCE Universidade Alto Vale do Rio do Peixe ( BRAZIL )".
- [14] Y. Heryanto, F. Farahinna, dan S. Wijanarko, "Evaluasi Responsivitas dan Akurasi : Perbandingan Kinerja ChatGPT dan Google BARD dalam Menjawab Pertanyaan seputar Python," vol. 9, hal. 248–256, 2024.
- [15] A. N. Aini, R. Rinofah, dan A. Maulida, "Efektifitas Pengembangan Artificial Intelligence (AI) pada Chatbot MbakPia," *JIIP - J. Ilm. Ilmu Pendidik.*, vol. 6, no. 10, hal. 8224–8228, 2023, doi: 10.54371/jiip.v6i10.3047.
- [16] B. Karyadi, "Pemanfaatan Kecerdasan Buatan Dalam Mendukung Pembelajaran Mandiri," *J. Teknol. Pendidik.*, vol. 8, no. 2, hal. 253–258, 2023, doi: 10.32832/educate.v8i02.14843.
- [17] S. Sugiono dan I. Agency, "Pemanfaatan Chatbot Pada Masa Pandemi Covid-19 Kajian Fenomena Society," no. November, 2022.
- [18] D. A. N. O. Pso, "INTI NUSA MANDIRI," vol. 18, no. 1, hal. 59–64, 2023.
- [19] B. Ram dan V. Pratima, "Artificial intelligence AI-based chatbot study of ChatGPT, Google AI Bard and Baidu AI," *World J. Adv. Eng. Technol. Sci.*, vol. 8, no. 1, hal. 258–261, 2023.
- [20] Carolus Borromeus Mulyatno, "Analisis User Experience Game Android : Amazepong Bali 1.0 Sigit," *J. Pendidik. dan Konseling Univ. Pahlawan Tuanku Tambusai*, vol. 4, hal. 1349–1358, 2022, [Daring]. Tersedia pada: <https://journal.universitaspahlawan.ac.id/index.php/jpdk/article/view/11165>
- [21] R. Umar, A. Z. Ifani, F. I. Ammatulloh, dan M. Anggriani, "Analisis Sistem Informasi Web Lsp Uad Menggunakan User Experience Questionnaire (Ueq)," *METHOMIKA J. Manaj. Inform. dan Komputerisasi Akunt.*, vol. 4, no. 2, hal. 173–178, 2021, doi: 10.46880/jmika.vol4no2.pp173-178.
- [22] M. Schrepp, "Buku Pegangan Kuesioner," *Marhaeni*, vol. 59, hal. 1–15, 2020.
- [23] M. S. Rahaman, M. M. T. Ahsan, N. Anjum, M. M. Rahman, dan M. N. Rahman, "The AI Race is on! Google's Bard and Openai's Chatgpt Head to Head: An Opinion Article," *SSRN Electron. J.*, 2023, doi: 10.2139/ssrn.4351785.
- [24] D. Khuntari, "Analisis Pengalaman Pengguna Aplikasi Gojek dan Grab dengan Pendekatan User Experience Questionnaire," *J. Tek. Inform. dan Sist. Inf.*, vol. 8, no. 1, hal. 275–286, 2022, doi: 10.28932/jutisi.v8i1.4499.
- [25] A. N. Hidayat, B. T. Hanggara, dan B. S. Prakoso, "Analisis Perbandingan Pengalaman Pengguna Pada Aplikasi Dompot Digital (Studi Kasus Pada OVO dan DANA)," ... *Teknol. Inf. dan Ilmu* ..., vol. 5, no. 6, hal. 2181–2187, 2021.
- [26] M. A. Kresnanto, B. T. Hanggara, dan B. S. Prakoso, "Analisis Pengalaman Pengguna pada Aplikasi Mobile Booking Hotel dengan menggunakan Metode User Experience Questionnaire ( UEQ ) ( Studi pada RedDoorz dan Airy )," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 10, hal. 3637–3646, 2020.
- [27] S. Rasio Henim dan R. Perdana Sari, "Evaluasi User Experience Sistem Informasi Akademik Mahasiswa pada Perguruan Tinggi Menggunakan User Experience Questionnaire," *J. Komput. Terap.*, vol. 6, no. 1, hal. 69–78, 2020, [Daring]. Tersedia pada: <https://jurnal.pcr.ac.id/index.php/jkt/>
- [28] E. Haerani dan A. Rahmatulloh, "Analisis User Experience Aplikasi Peduli Lindungi untuk Menunjang Proses Bisnis Berkelanjutan," *SATIN - Sains dan Teknol. Inf.*, vol. 7, no. 2, hal. 01–10, 2021, doi: 10.33372/stn.v7i1.762.
- [29] A. Noor dan E. L. Hadisaputro, "Analisis Pengalaman Pengguna Pada Aplikasi TIX ID Menggunakan Metode User Experience Questionnaire," *J. Inf. Syst. Res.*, vol. 3, no. 4, hal. 672–677, 2022, doi: 10.47065/josh.v3i4.1881.
- [30] R. A. Ningsih dan K. D. Tania, "Comparison of Tiket.com and Pegipegi User Experience Using the UEQ Method," *J. Tek. Inform. dan Sist. Inf.*, vol. 9, no. 2, hal. 177–188, 2023, doi: 10.28932/jutisi.v9i2.6079.