

ARE ALARM BELLS RINGING IN ACADEMIA? CHATGPT AS A SAMPLE OF USING CHATBOTS IN EDUCATION

Dr. Gurhan DURAK

ORCID: 0000-0003-2944-3713
Faculty of Education
Balıkesir University
Balıkesir, TURKIYE

Dr. Serkan CANKAYA

ORCID: 0000-0002-3951-9809
Faculty of Economics and Administrative Sciences
Izmir Democracy University
Izmir, TURKIYE

Received: 01/09/2023 **Accepted:** 28/11/2023

ABSTRACT

The main aim of this mixed methods study is to determine the evaluations made by academics on an academic text generated with ChatGPT and get their general views on ChatGPT. The convergent parallel design is utilized where the qualitative and quantitative methods are combined to produce triangulated results. Data were collected via an online form from 45 academics. Academics were asked to read and rate the sample text generated by ChatGPT. The titles "Writing rules," "Language and expression," and "Subject integrity" of the text generated by ChatGPT received the highest ratings from the academics. The qualitative findings obtained were interpreted within the framework of the "Diffusion of Innovations" and "Uses and Gratifications" theories, and the related themes were formed. As a result, the participants were asked about their feelings about using ChatGPT, and the themes of astonishment, pessimism, anxiety, and hopefulness emerged. Participants were asked their opinions about the benefits and negative aspects of using ChatGPT. Themes emerged under the benefits, and negative aspects titles were evaluated. Finally, the other themes that emerged under the satisfaction with ChatGPT and making ChatGPT widespread titles were assessed.

Keywords: Artificial intelligence, Chatbots in Education, ChatGPT, Diffusion of Innovations, Generative AI.

INTRODUCTION

Over the years, many chatbots have been released, like ELIZA in 1966, PARRY in 1972, Jabberwacky in 1988, Dr Sbaitsio in 1992, ALICE in 1995, SmarterChild in 2001, Siri in 2010, Google Assistant in 2012, Cortana in 2014, and Alexa in 2014. Although chatbots have been around for a while, users have only begun to notice them in recent years. This shift in recognition of chatbots and conversational interfaces mainly resulted from advancements in AI and machine learning and the rising use of messaging applications (Ina, 2022). ChatGPT is one of the chatbots launched by OpenAI on November 30, 2022. In a short time, it has gained tremendous popularity by being used by many people (Haleem et al., 2022).

G.P.T. ("Generative Pre-training Transformer") is a machine learning model developed by OpenAI. It is an extensive, deep neural network trained to generate natural language text that is coherent and similar to human writing. GPT models have achieved impressive results on various natural language processing tasks, including language translation, summarization, and question answering. It can generate human-like text and perform a wide range of language tasks. It is designed to understand and respond to natural language input and provide helpful and accurate information to users. ChatGPT is a sibling model to InstructGP and was fine-tuned from a model in the GPT-3.5 series, which finished training in early 2022 (OpenAI, 2022).

OpenAI’s ChatGPT-3 is a state-of-the-art language processing AI model. One of the most extensive and advanced language processing AI models, it has 175 billion parameters. In addition to answering questions, composing essays and poetry, summarizing lengthy books, translating languages, conversing convincingly, and even creating computer code, GPT -3 can produce anything with a linguistic structure. Online text datasets were used to train the model. Over 45 TB of unfiltered text and 570 GB of filtered content from books, web texts, Wikipedia articles, and other online writings are included in the training. Three hundred billion words were entered into the system (Kumar, 2022). GPT-4, the most recent version of OpenAI’s language model system, was formally released on March 13, 2023. Users may use the Chat GPT-4 tool with a premium membership. According to OpenAI, GPT-4 is more advanced in creativity, visual input, and longer context and surpasses ChatGPT in its advanced reasoning capabilities (OpenAI, 2023).

Using the same techniques as InstructGPT, ChatGPT was trained using Reinforcement Learning from Human Feedback (RLHF), with a few minor variations in the data configuration. Human AI trainers acted as users and AI assistants and provided conversations. Then, they ranked randomly selected model-written messages and sampled several alternative completions (OpenAI, 2022).

ChatGPT rarely can provide responses that are erroneous or illogical because(1) there is no source of truth during Reinforcement Learning training; (2) making the model more cautious makes it decline questions that it can answer correctly; (3) supervised training deceives the model. The model frequently can employ certain words and phrases excessively. These problems are due to over-optimization issues and biases in the training data (trainers favour lengthier replies that appear more comprehensive). When the user provides an uncertain query, the model should offer to clarify questions. Instead, the present model typically infers what the user means (OpenAI, 2022).

With the announcement of ChatGPT on November 30, 2022, ChatGPT gained people’s attention immediately. Through social media, people talked about ChatGPT and shared the texts generated. Meanwhile, Figure 1 shows ChatGPT searches via Google Trends. With its announcement at the end of November, it has reached peak popularity (100%) in a very short time, which has caused quite a significant impact worldwide. The values indicate how many searches have been conducted for that particular term compared to the total number of searches performed on Google (Google News Initiative, 2019).

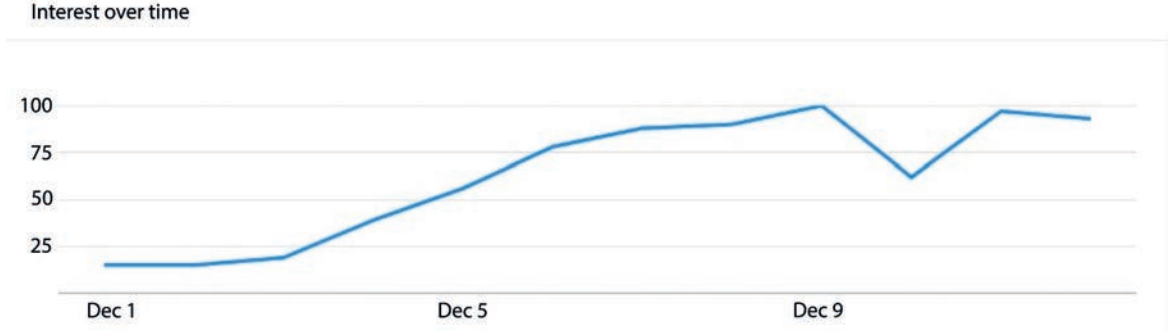


Figure 1. ChatGPT on Google Trends

The “interest by region” option has been selected to determine which regions this intense demand is high. According to Figure 2, this rate is 100% in China. Norway, Finland, Portugal, and Israel follow China.

Interest by region



Figure 2. Search Interests by Region About ChatGPT

Besides other skills, ChatGPT can be successful in academic writing. A good example is how ChatGPT can write an article's Introduction and Literature Review sections. It can enrich the text using citations, just like an academic article. On the other hand, it may not always provide accurate citations for the sources it uses to generate text (Donmez et al., 2023). Articles generated with ChatGPT will likely appear soon. This situation will bring about a change that concerns academic journals, editors, referees, and academicians. While ChatGPT can be helpful in academic writing, it can cause various concerns, such as ethics. Most people who publish, edit, and referee in academic journals are academics. So, ChatGPT's abilities, which is a new technology, in writing articles must be evaluated from various academic perspectives. As ChatGPT technology is relatively new, we have yet to find research evaluating the academic writing capabilities of ChatGPT through the lenses of academics. This study will fill an essential literature gap and pioneer the field. In this context, this research aims to determine the evaluations of academics about an academic text generated with ChatGPT and to get their opinions about ChatGPT. In this direction, the following research questions were addressed.

- Our study was designed to answer the following research question in the quantitative phase: How do academics rate the academic text generated by ChatGPT?
- Our study was designed to answer the following research question in the qualitative phase: What are the opinions of academics who have used ChatGPT about the positive/negative aspects of ChatGPT and possible use in the future?

LITERATURE REVIEW

In the literature, it is seen that there are many studies on chatgpt and other chatbots. These studies have generally focused on the possible benefits and threats of their use in education. To our knowledge, our work is the first study that qualitatively analyzed academics' sentiments and feedback on ChatGPT's academic writing capabilities. We contribute to the literature by providing a snapshot of the early public responses to this latest technology.

In a study (Adiguzel et al., 2023), possible uses of ChatGPT in the field of education were examined. According to the study, while modern chatbots like ChatGPT offer numerous benefits, their implementation in education raises significant ethical and practical concerns. The authors aim to offer valuable insights into how AI can be effectively integrated into education to aid both teachers and students, while also encouraging responsible and ethical use. In a similar study, Halaweh (2023) offered real-world examples of how ChatGPT may be used for academic writing. The author suggested a policy for universities and instructors about ChatGPT to match the specific demands of their institutions and courses.

Susnjak (2022) carried out a research using ChatGPT to evaluate its capacity for critical thinking as opposed to just knowledge retrieval. High levels of precision, coherence, and accuracy were shown in the results. Similar to this, Khalil and Er (2023) tested how well ChatGPT-produced writings might be identified by plagiarism detection software. Out of the 50 essays examined, 40 had a similarity score of 20% or less, suggesting a high degree of originality, according to their findings. Contrarily, Dowling and Lucey (2023)

noted that while ChatGPT has strengths in idea generation and data identification, it has problems in literature synthesis and the creation of useful testing frameworks in the field of finance research.

In another study, the researchers employed a three-stage instrumental case study approach, which involved analyzing tweets using social network analysis, conducting interviews and analyzing their content, and investigating the experiences of users, all to explore the concerns associated with using chatbots in education, particularly ChatGPT, among early adopters. The findings of the study indicated that while ChatGPT is a potent tool for educational purposes, it should be used with caution, and more guidelines on its safe use in education need to be established. Additionally, the study identified several research areas and questions that researchers and practitioners should explore for a safer and better adoption of chatbots, particularly ChatGPT (Tlili et al., 2023).

A meta-analysis study investigated the effectiveness of chatbot technology in achieving educational outcomes. The results indicated that chatbots had a medium-to-high overall effect size on educational outcomes, regardless of intervention duration, chatbot roles, and learning content. Chatbot technology had a positive impact on explicit reasoning, learning achievement, knowledge retention, and learning interest. However, chatbots did not significantly improve critical thinking, learning engagement, and motivation. These findings suggest that teachers and instructors can use appropriate teaching methods to facilitate sustainable education based on the benefits that chatbots offer (Deng & Yu, 2023).

Chocarro et al. (2023) investigated the teachers' usage intention of a chatbot in an educational setting, taking into account conversational design decisions and teachers' characteristics. The study found that teachers' perceived usefulness of the chatbot had a significant and positive impact on their intention to use the technology. It was observed that enhancing the performance and usefulness of chatbots is crucial in encouraging teachers to adopt this technology for their professional duties.

Sok and Heng (2023) examined the advantages and disadvantages of utilizing ChatGPT in the fields of education and research. They suggested that ChatGPT has five primary benefits, which include facilitating learning assessments, improving teaching techniques, providing virtual personal tutoring, generating an outline, and encouraging the development of ideas. Nonetheless, they indicated that there are also certain risks associated with using ChatGPT, such as academic dishonesty, unfair evaluation of learning, inaccurate data, and an excessive dependence on artificial intelligence.

Among the studies in the literature, the most similar study to our study was conducted by Donmez et al. (2023). They investigated the advantages and difficulties encountered when using the ChatGPT application in scientific research. They indicated that although artificial intelligence technologies offer advantages to researchers in terms of validation, innovation, and presenting various perspectives, there are issues with reliability when producing content and ethical concerns regarding plagiarism. They concluded that artificial intelligence does not write articles for researchers but rather provides suggestions and assistance during the article-writing process.

In a study on the use of ChatGPT at the university level, data were collected from 7 academics and 14 PhD students and the responses were analyzed using thematic content analysis. According to this study the themes emerged as "Evolution of learning and education systems", "changing role of educators", "impact on assessment and evaluation", "ethical and social considerations", "future of work and employability", "personalized learning", "digital literacy and AI integration", "AI as an extension of the human brain", and "importance of human characteristics".

THEORETICAL FRAMEWORK

This study is designed within the diffusion of innovations theory (DIT) and Uses and Gratifications theories (UGT). It can be said that these theories are generally used in studies on communication, media and the use of new technologies. In one study, McQuail and Windahl (1993) stated that, from a UGT perspective, researchers should seek answers to the questions "Why do people use media?" and "What do they use it for?" rather than "how does media affect people?" Within the UGT framework, it is also possible to explain why people use new technologies such as ChatGPT and for what purpose (Durak, 2017).

According to DIT by Rogers(1995), innovation can be an idea, an application, or an object that the individual or organization considers novel. Ultimately, innovation need not be an unknown concept or design because if the individual or organization has not yet utilized it, it can be considered innovative (Berger, 2005). Innovation can be defined theoretically as a new product, technology, perspective, or solution to a problem for individuals or organizations (Demir, 2006). Rogers (1995) noted that, according to the model, the innovation decision-making process consisted of five phases: knowledge, persuasion, decision, adoption, and confirmation. During the information phase, the individual receives details about innovation and its advantages. Next, in the convincing phase, the individual evaluates this innovation's negative and positive aspects and adjusts their attitude accordingly. In the decision phase, the individual accepts or rejects the innovation. If the individual accepts, the application phase will occur. In the approval phase, the individual confirms and strengthens the fitness-related decision (Orr, 2003). In this study, ChatGPT is regarded as an innovation for academics because they did not use it previously.

METHOD

This section covers the research model, data collection tools, study participants, data analysis, validity and reliability studies and the application stage of the study.

Research Model

In this study, the convergent parallel design, one of the mixed method designs, is utilized. This design combines qualitative and quantitative methods to produce triangulated results. Two data sets are initially collected simultaneously and then analyzed using quantitative and qualitative analytical techniques (Creswell & Plano Clark, 2018; Shorten & Smith, 2017). In a convergent design, combining quantitative and qualitative data will enable the researcher to comprehensively understand the one provided by quantitative or qualitative results alone. It is a method that combines two data sets to obtain a complete picture of the issue being investigated and to validate one set of findings with the other (Creswell & Plano Clark, 2018).

Data Collection Tools

In this study, we employed a mixed-methods approach, collecting both quantitative and qualitative data through an online form designed by the researchers. The form included questions aimed at gathering quantitative data, such as demographic information of the participants, their publication count, and the average number of manuscripts they review annually. Additionally, we incorporated a 5-point Likert scale for academics to quantitatively evaluate the academic text generated by ChatGPT. This scale was meticulously developed by analyzing manuscript evaluation criteria commonly used in major journals, covering aspects like originality, integrity, qualification, use of resources, problem identification, purpose, significance, adherence to academic writing standards, and language/expression. Qualitative data were also gathered through the form, focusing on participants' detailed opinions and evaluations about ChatGPT. This part of the data collection aimed to provide a deeper understanding of their perspectives and insights, which goes beyond the quantifiable aspects. The development of the interview questions was informed by the theories underpinning our study, ensuring that both qualitative and quantitative data were aligned with our research objectives. To enhance the tool's validity and reliability, educational technology experts specializing in measurement and evaluation reviewed the form. Their feedback was instrumental in refining and finalizing the tool, ensuring it effectively captured both quantitative and qualitative dimensions of the participants' responses.

Participants

Within the scope of the study, 45 academics from different universities were reached with convenience sampling methodology, and the data collection tool was applied. Since this study focuses on the use, advantages, and disadvantages of chatbots in academic publishing, the participant group was determined as academics by nature. The participants were categorized according to the fields of expertise stated in Table 1.

Table 1. Participants and their field of expertise

| Group | Field of Expertise | Number of Participants |
|-------|--|------------------------|
| 1 | Computer Education and Instructional Technologies (CEIT) | 22 |
| 2 | Distance Education (DE) | 13 |
| 3 | Other Departments (OD) | 10 |

According to Table 1, nearly half of the participants are academics from the CEIT department, while a significant portion is from DE, and the rest are from OD. Participants are coded with fields of expertise like CEIT -1, DE-7, and OD-4.

Data Analysis

The data collection tool, prepared within the scope of theoretical foundations in the study, was delivered to participants via Google forms through the researchers' personal networks. Quantitative data were analyzed in spreadsheet software, and graphics were created to represent the data. In the interview form, the participants were given detailed information about the purpose and method of the research. The questions and responses to the questions were transferred into a spreadsheet table under the prepared coding draft.

The academics' responses to the questions were examined one by one by two researchers, and each researcher formed his own interview coding key. A field expert evaluated the consistency of the coding keys, and the reliability study of the coding key was carried out. After concluding that consistency was achieved, the themes were formed within the scope of the study's theoretical framework. In addition, direct quotations are frequently used to effectively reflect academics' general views.

Validity and Reliability

The participants were informed that their views would be used only for academic purposes and that their names would be confidential. With these precautions, it was thought that they could express their opinions freely. The coding keys generated by the researchers were compared regarding reliability and consistency. Intercoder reliability is a measure of the consistency or agreement between two or more coders who are independently coding the same qualitative data (O'Connor & Joffe, 2020). It is an important parameter in qualitative research as it helps to ensure the validity and reliability of the findings. Cohen's Kappa is a statistical measure of inter-rater agreement that takes into account the possibility of chance agreement (Rau & Shih, 2021). It ranges from -1 to 1, with 1 indicating perfect agreement. The Cohen Kappa value between the two coders was calculated as 0.92 indicating almost perfect agreement. Afterwards, themes were created according to the coding keys and theoretical framework. Besides, direct quotations were included to reflect the views of academics truthfully and straightforwardly. In addition, the responses of academics were stored in a safe place to allow other researchers to review them on demand.

Application Stage

In this study, a total of 45 academics from different universities were reached, and the ChatGPT software was briefly introduced to these academics with various sample prompts. Then they were asked to test the ChatGPT software themselves for a week. The following command was entered in ChatGPT: "Write the literature review part of the article Investigating Collaborative Flipped Learning in approximately 1000 words". ChatGPT generated a 400-word text on the subject after this command (Appendix 1). The reasons for choosing flipped learning as a topic are that it is a popular method that has been widely studied and that it is close to the participants' fields of study. The text generated by ChatGPT was added to the data collection tool, making no additions. ChatGPT can generate a different text each time even if the same prompt is used. Afterwards, the online data collection tool was applied, and the academics were asked to read this text and rate it in terms of various titles.

Limitations of the Study

Although the participation group of this study is sufficient for qualitative methods, the fact that a higher number of academics were not reached for the analysis of quantitative data can be considered as a limitation. This sample size may not be sufficient to generalize the findings to a broader academic community. The concentration of participants from certain fields may limit the diversity of perspectives and insights into the broader applicability of chatbots in various academic disciplines. The use of an online form for data collection, although practical, may limit the depth of responses. In-person interviews or focus groups could provide more nuanced insights. The text generated by ChatGPT, being limited to 400 words, may not have fully captured the complexity or depth of the topic, potentially influencing the academics' evaluations.

FINDINGS

Within the scope of the study's quantitative data, the figures representing demographic data, professional experiences, the number of published articles and the annual average number of manuscripts reviewed will be presented. In addition, the participants' evaluations regarding the sample text generated by ChatGPT will be interpreted. As for the qualitative part of the study, the academics' responses were analyzed and coded, and the themes were generated from the coded data. The qualitative findings obtained were interpreted according to both DIT and UGT and were grouped according to (1) general feelings about ChatGPT, (2) The benefits and negative aspects of using ChatGPT and (3) the use and spread of ChatGPT.

General Information about Academics

In Figure 3, there is a graph of the experiences of 45 academicians in terms of working years. According to the graph, more than half of academics have a professional experience of 11-20 years.



Figure 3. Professional Experience of the Academics

The data on the number of articles published by the participants and the annual average number of manuscripts reviewed by the participants are supplied in Figure 4.

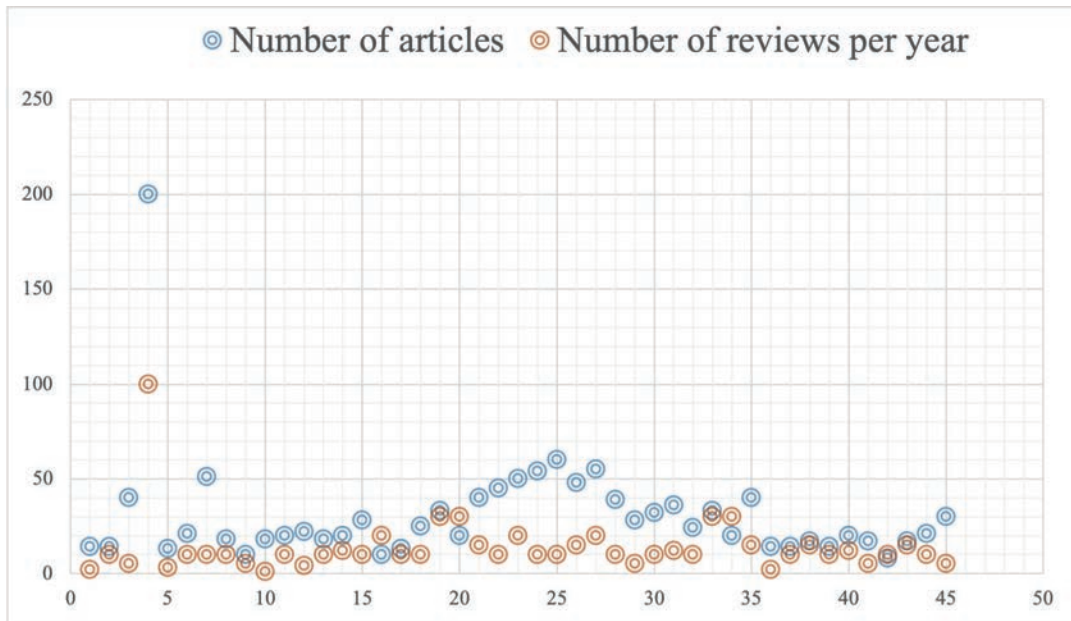


Figure 4. The Number of Articles and Annual Average Number of Manuscripts Reviewed by The Participants

According to Figure 4, it is found that the participants have an average of 10 to 60 articles (one academicians with 200 articles is considered an exception), and the annual average number of manuscripts reviewed by the participants is between 1 and 30 (100 reviews are considered an exception). In general, it can be said that the number of articles published by the participants is in a good ratio according to their professional experience. Similarly, it can be thought that the annual average number of manuscripts reviewed by the participants is reasonable.

The participants were asked to rate the text produced by ChatGPT on a scale ranging from very poor to very good based on the various titles. Figure 5 displays the results of 45 participants' ratings of the text generated by ChatGPT.

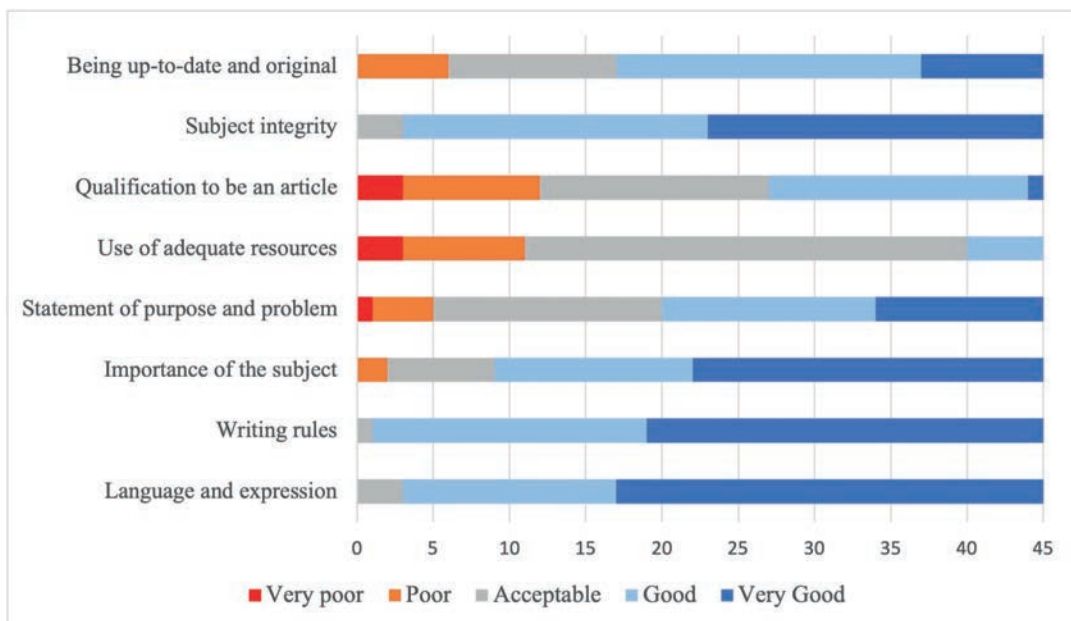


Figure 5. The Evaluation Form by the Participants of the Text Produced by ChatGPT

According to Figure 5, in the evaluation of the text created by ChatGPT, the titles that received the highest ratings (good + very good) from the participants were “Writing rules”, “Language and expression”, and “Subject integrity”. These titles are followed by “the importance of the subject”. “Use of adequate resources” and “qualification to be an article” are among the topics that receive low participant ratings. The titles “Being up-to-date and original” and “statement of purpose and problem” received different ratings from the participants.

General Feelings about ChatGPT

The participants were asked about their feelings about using ChatGPT. The findings were reduced into themes, and their frequencies are supplied in Table 2.

Table 2. General Feelings About Using ChatGPT

| Theme | Frequency (f) |
|-----------------|---------------|
| General feeling | |
| Astonishment | 19 |
| Pessimism | 13 |
| Anxiety | 9 |
| Hopefulness | 7 |

It was found that most participants stated that they felt astonished after the ChatGPT experience. This feeling is followed by Pessimism, Anxiety and Hopefulness. About the “astonishment” theme, CEIT -7 expressed her thoughts as follows: “*I was quite amazed. This is a decent introduction. A few graduate students can compose a text like this.*” As for the Pessimism theme, DE-2 stated, “*It is obviously a terrible situation that people can write articles without any effort,*” whereas CEIT -3 stated, “*There is no reason why a graduate student who speaks a foreign language cannot write homework and article proposals with ChatGPT. They can graduate easily...*” While it was observed that some participants were hopeful about the development of such software, one participant, OD-1, said, “*This version has been very successful, and I am very hopeful about future versions.*”

The Benefits and Negative Aspects of Using ChatGPT

The findings and frequencies regarding the participants’ opinions about the benefits and negative aspects of using ChatGPT are presented in Table 3.

Table 3. The Benefits and Negative Aspects of Using ChatGPT

| Theme | Frequency (f) |
|--|---------------|
| Benefits | |
| Time & energy saving | 29 |
| Contribution to academic writing | 19 |
| Facilitating role | 15 |
| An innovative software | 13 |
| Increase in the number of publications | 9 |
| Negative aspects | |
| Individual effects | |
| Negative effects on reading habits | 24 |
| Reduces originality | 19 |
| Negative effects on creativity | 14 |
| Laziness | 10 |
| Academic effects | |
| Academic dishonesty | 21 |
| May cause monotonous writing | 18 |
| Fake publications will increase | 16 |
| May terminate some sectors | 12 |

The themes of benefits and negative aspects emerged from the answers given by the participants. The theme of negative aspects is also divided into individual and academic effects. The most prominent title under the theme of the benefits is “time & energy saving”, followed by “contribution to academic writing” and “facilitating role”. Regarding the “time & energy saving” theme, which most participants emphasized, CEIT-5 stated that *“I think this software will save time in terms of the introduction and literature review sections, on which a lot of time and effort is spent”*, while OD-2 commented: *“It will make things easier and significantly reduce the effort spent in the introduction sections, which are difficult places to write”*. DE-5 stated, *“I think it will facilitate the academic writing processes and make the processes more efficient”* on the theme of “contribution to academic writing”, and OD-7 stated, *“I think it can make it much easier to write introduction part in articles”*. Regarding the facilitating role theme, CEIT-16 said, *“it can be used to help academics, especially in literature review”*, while CEIT-21 said, *“Content can be created for the introduction section. The software can make serious contributions to the literature review”*.

The responses given under the theme of negative aspects were evaluated as individual and academic effects. As for the theme of “individual effects”, the factors of “negative effects on reading habits” and “reduces originality” come into prominence. These views are followed by “negative effects on creativity” and “Laziness”. Regarding the theme of reading habits, CEIT-10 stated that *“It may have a negative impact on individuals’ reading habits and their ability to investigate the subject”*, while OD-2 concurred and stated, *“It may diminish research and reading skills.”* Regarding the theme of “reduces originality”, many participants stated that the ChatGPT-generated texts would decrease the originality in publications. CEIT-9 and OD-1 defended the same view and commented: *“Eventually, there will be problems with the authenticity of the studies...”*. Almost all 10 participants who put forward their view regarding the theme of laziness expressed with similar expressions that such software can cause laziness after a certain period.

The “Academic effects” theme under the “negative aspects” theme includes the academic evaluation of negative situations related to ChatGPT. Among the answers given by the participants, the most prominent are “academic dishonesty” and “may cause monotonous writing”. These themes are followed by “fake publications will increase” and “may terminate some sectors”. CEIT-6, one of the many participants who expressed their opinions on the theme of academic dishonesty, made the following comment: *“This software can handle most of the article writing. The author can submit the article to journals after a proofreading process. This will lead to ethical violations”*. A participant of DE-4 stated his the title “may cause to monotonous writing” as *“... it may cause to monotonous writing because the articles will come out of the same environment”*. As for the theme of “fake publications”, which many participants demonstrated, DE-7 expressed his concern *“I am afraid that fabricated articles will be written using fake data with this software”*. Among the answers given, one of the most negative ones is under the “may terminate some sectors” theme. CEIT-13, one of the participants who advocated this view, said, *“ChatGPT may cause the existence of the academic profession to disappear.”* Another participant, OD-5 expressed his opinion that *“if the writing style is differentiated by ChatGPT, then academic activities such as academic writing, publishing articles and presenting papers may come to an end”*.

Use of Spread of ChatGPT

The findings and frequencies regarding the participants’ experience and views of ChatGPT and their opinions about making ChatGPT widespread can be seen in Table 4.

Table 4. Findings Regarding the Use and Spread of ChatGPT

| Theme | Frequency (f) |
|---|---------------|
| Satisfaction with ChatGPT | |
| Positive | |
| ChatGPT can be used as a guide for academic writing. | 18 |
| The text is well organized | 11 |
| ChatGPT explained the subject very well | 7 |
| Negative | |
| The sample text is insufficient to be an academic article | 7 |
| Neutral | |
| ChatGPT can be used for simple tasks like homework etc. | 9 |
| Making ChatGPT widespread | |
| Positive | |
| ChatGPT can be used as supplementary | 24 |
| The text generated by ChatGPT can be used as a draft | 15 |
| Neutral | |
| ChatGPT can be used in non-academic studies | 16 |
| ChatGPT can be used after checking for plagiarism | 9 |
| Negative | |
| ChatGPT should never be used in academic publications | 9 |
| There will be a problem of originality in publications | 8 |

The participants were asked their opinions about their satisfaction with ChatGPT and making ChatGPT widespread. The themes under satisfaction with ChatGPT are classified as positive, neutral and negative. Regarding the theme of positive satisfaction, “ChatGPT can be used as a guide” has become prominent among responses. This view is followed by “the text is well organized” and “ChatGPT explained the subject very well” themes. CEIT -8 expressed his thoughts: *“I intend to use ChatGPT in the future. An assistant, who can scan the literature faster and more accurately than I can, will make my work easier as it is a guide”*. OD-4, one of the many participants who expressed a positive opinion about the “the text is well organized” theme, said, *“I was really surprised when I examined the sample text. The transition between paragraphs is very good. It has prepared a very fluent text”*. DE-7 said, *“I will use ChatGPT. It is very successful, especially in organization and giving ideas”*. Although most participants had positive opinions about ChatGPT, there were also some negative views. Participants who think this way generally stated that the sample text is insufficient to be an academic article. CEIT-11 said that *“ChatGPT has created a type of text that can be considered successful, but not every text created can be an academic publication. The limited use of resources and the reliability issues make me think seriously”*. Another participant stated that he was partially satisfied with the text created by ChatGPT and commented that *“it can be used in simple homework, although not in academic studies.”*

Finally, the theme of “making ChatGPT widespread” is classified as positive, neutral and negative. The themes “can be used as supplementary” and “can be used as a draft” emerged under the positive theme. Within the scope of positive opinions, CEIT-17 commented that *“ChatGPT can be used as a software to help researchers”*, while DE-6 stated that *“each researcher can have an AI assistant that helps her train according to his/her own rules”*. Under the theme “The text generated by ChatGPT can be used as a draft”, CEIT-14 commented that *“text produced by ChatGPT can be considered as a draft”*. Under the negative theme, some participants stated that ChatGPT should not be used in academic studies. Some participants emphasized that there may be problems with the originality of the publications. DE-7 stated that *“this kind of software should never be used in academic publications. The resulting text has no credibility. It is unclear where it was taken from. It will also bring ethical problems.”* As for the originality concern in publications, CEIT-19 stated that *“after a certain period of time, the same type of publications may be encountered. However, the human factor is very important, writing styles are important... Without these, academic publications will consist of repetitive articles.”*

CONCLUSION, DISCUSSIONS AND SUGGESTIONS

When the evaluations of the academics about the academic text generated with ChatGPT are examined, quite surprising results are obtained. It has been observed that the sample text received very successful evaluations from academics in the title of “writing rules” and “language and expression”. Qualitative data also support this result. Most academics have argued that the text generated by ChatGPT is quite successful. On the other hand, academics considered that the references in the sample text were not up-to-date and that not enough references were used as a deficiency. Some academics’ opinions that the generated text cannot be an academic publication on its own support this situation.

Due to the open-ended questions, it was revealed that the academics tried such a software for the first time, and it was seen that they mostly felt “astonishment”. It can be expected that people will be surprised when they see what artificial intelligence can do. Still, the emergence of emotions such as pessimism and anxiety besides astonishment needs to be considered. Some academics stated that they had hesitations about ChatGPT and criticized the use of ChatGPT in academic publications.

Among the positive opinions that emerged under the theme of The Benefits and Negative Aspects of Using ChatGPT, it was stated that the theme of “time & energy saving” was expressed by most academics and that many stated that ChatGPT could be used as a facilitator and a contributor in academic writing. Speed and time saving were expressed as a core benefit of ChatGPT in literature (Eke, 2023; Fido & Wallace, 2023; Rathore, 2023). In line with these views, this software is seen by academics as a qualified and successful software in academic writing. Considering that with ChatGPT, meaningful texts can be created in a very short time on any academic topic; plus, it can be thought that such views of academics are as expected. Considering the opinions of academics about ChatGPT within the framework of uses and gratifications theory, it is seen that satisfactions such as “time & energy saving with ChatGPT”, “facilitating role of ChatGPT”, “ChatGPT can be used as a supplementary tool”, among the positive opinions that have emerged, are also included in various studies in the literature (Dunne et al., 2010; Durak, 2017; Lim & Ting, 2012; Tinmaz, 2011; Wang et al., 2018; Wodzicki et al., 2012). Similarly, Donmez et al. (2023) stated that the texts produced by chatbots cannot be used directly in article writing, and that it is more appropriate to use such software as a support.

There have been negative opinions about ChatGPT, such as that it may negatively affect reading habits, reduce the originality of publications, and negatively impact the creativity of academics. It is normal that some people have negative thoughts towards technological innovations. Some academics have expressed the view that ChatGPT can cause academics to be lazy. The fact that ChatGPT significantly facilitates academic writing for academics may lead to this result. Other themes that emerged about the negative effects of ChatGPT are “academic dishonesty”, “may cause monotonous writing”, “fake publications will increase”, and “may terminate some sectors”. Similarly, there are some studies that raise ethical concerns about the use of ChatGPT in education as misinformation generation, bias and privacy, responsible implementation and leadership to ensure the ethical use of AI in education, and serious consequences of using ChatGPT in education and science (Baidoo-anu & Ansah, 2023; Firat, 2023; Willems, 2023). One of the themes that should be emphasized here is “may terminate some sectors”. The emergence of technological innovations brings concerns about the future of many sectors. It is one of the assumptions spoken today that the developments in robotics and artificial intelligence fields will change professions such as medicine, the military, etc., in the future.

Finally, when the academics were asked whether they would use ChatGPT in the future, it was revealed that most thought of doing so. Sharing texts generated with ChatGPT by some academics on social media platforms supports this result. The theory of diffusion of innovations can explain this situation. Within the framework of this theory, the fact that academics have no previous ChatGPT experience means this application is an innovation for them. From the findings, despite serious concerns, most academics thought that ChatGPT would be useful in the academic field and stated positive opinions about it, showing they adopted this innovation. Considering that the academics thought the spread of ChatGPT, new versions of ChatGPT would benefit the academic community can be explained by a diffusion of innovations.

According to the results of this research and considering the advantages /disadvantages of using ChatGPT, these recommendations were made:

- Instead of a short text generated with ChatGPT, like in this study, a more comprehensive text, including other article sections, can be generated with ChatGPT, and qualitative and quantitative studies can be carried out on that text.
- Interdisciplinary research projects can be encouraged where experts from fields like computer science, linguistics, education, and psychology collaborate to explore the implications of ChatGPT and similar technologies in academic writing and research.
- It would be beneficial to conduct longitudinal studies to observe how the use of ChatGPT and similar AI tools evolve over time in academic settings. This would provide insights into the long-term impacts of these tools on academic writing, originality, and creativity.
- There is a need to develop comprehensive ethical guidelines and policies for using AI tools in academic writing. These should address concerns about academic dishonesty, originality, and the potential for AI to create monotonous or generic content.
- Integrating ChatGPT into academic curricula as a supplementary tool, teaching students how to effectively and ethically use such AI technologies in research and writing can be considered.
- AI literacy programs can be implemented for academics and students to enhance their understanding of AI capabilities, limitations, and ethical implications. This would help in informed and responsible use of AI tools in academia.
- Comparative studies can be performed to assess the performance of ChatGPT against other AI writing tools in academic settings. This can provide a clearer understanding of the strengths and weaknesses of various AI technologies in academic applications.
- The impact of AI tools like ChatGPT can be assessed on different academic professions, particularly in terms of job roles, responsibilities, and the evolving nature of academic research and publication.
- The effectiveness of ChatGPT and similar tools can be explored in different cultural and linguistic contexts. This is vital for understanding how AI tools perform in diverse academic environments.
- Public perception and acceptance of AI tools in academic publishing can be investigated. Understanding societal attitudes towards AI in academia can guide more responsible and community-accepted applications.

BIODATA and CONTACT ADDRESSES of AUTHORS



Dr. Gurhan DURAK is a Professor of Instructional Technology at Balikesir University. Dr. Durak gained his Ph.D. in Distance Education at Anadolu University. His academic interest areas are artificial intelligence(ai) in education, generative ai, open and distance learning, virtual reality, metaverse in education, social learning networks, support services in distance education, online learning, and administration of distance education. He has authored over 50 articles, book contributions, books, and editorials, and is a Chief Editor of the Journal of Educational Technology and Online Learning, as well as a member of the Editorial Boards of other international journals.

Gurhan DURAK
Instructional Technologies, Faculty of Education
Address: Balikesir University, 10100, Balikesir, Turkiye
E-mail: gurhandurak@balikesir.edu.tr



Dr. Serkan CANKAYA is Associate Prof. Dr. in the Department of Management Information Systems at Izmir Democracy University. Dr. Cankaya received an M.S. degree in the Computer Education & Instructional Technology Department at Balikesir University. He gained his Ph.D. in Computer Education & Instructional Technology Department at Anadolu University at June 2013. His academic interest areas are artificial intelligence(ai) in education, generative ai, open and distance learning, virtual reality, metaverse in education, social learning networks, support services in distance education, online learning, and administration of distance education. He has authored over 50 articles, book contributions, books, and editorials, and is a Chief

Editor of the International Journal of Computers in Education, as well as a member of the Editorial Boards of other international journals.

Serkan CANKAYA

Management Information Systems, Faculty of Economics and Administrative Sciences

Address: Izmir Democracy University, 35140, Izmir, Turkiye

E-mail: serkancankaya79@gmail.com

REFERENCES

- Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI : Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology, 15*(3).
- Baidoo-anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *SSRN*.
- Berger, J. I. (2005). Perceived consequences of adopting the Internet into adult literacy and basic education classrooms. *Adult Basic Education, 15*(2), 103–121.
- Chocarro, R., Cortinas, M., & Marcos-Matas, G. (2023). Teachers' attitudes towards chatbots in education: a technology acceptance model approach considering the effect of social language, bot proactiveness, and users' characteristics. *Educational Studies, 49*(2), 295–313. <https://doi.org/10.1080/03055698.2020.1850426>
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). Sage.
- Demir, K. (2006). Rogers'in yeniligin yayilmasi teorisi ve Internetten ders kaydi. *Egitim Yonetimi Dergisi, 12*(47), 367–391.
- Deng, X., & Yu, Z. (2023). A Meta-Analysis and Systematic Review of the Effect of Chatbot Technology Use in Sustainable Education. *Sustainability, 15*(4). <https://doi.org/10.3390/su15042940>
- Donmez, I., Idil, S., & Gulen, S. (2023). Conducting academic research with the AI interface ChatGPT : Challenges and opportunities. *Journal of STEAM Education Journal, 6*(2), 101–118. <https://doi.org/https://doi.org/10.55290/steam.1263404>
- Dowling, M., & Lucey, B. (2023). ChatGPT for (Finance) research: The Bananarama Conjecture. *Finance Research Letters, 53*(January), 103662. <https://doi.org/10.1016/j.frl.2023.103662>
- Donmez, I., Idil, S. & Gulen, S. (2023). Conducting academic research with the AI interface ChatGPT: Challenges and opportunities. *Journal of STEAM Education, 6*(2), 101-118. <https://doi.org/10.55290/steam.1263404>
- Dunne, A., Lawlor, M. A., & Rowley, J. (2010). Young people's use of online social networking sites - a uses and gratifications perspective. *Journal of Research in Interactive Marketing, 4*(1), 46–58. <https://doi.org/10.1108/17505931011033551>
- Durak, G. (2017). Using social learning networks (SLNs) in higher education: Edmodo through the lenses of academics. *International Review of Research in Open and Distributed Learning, 18*(1), 84–109. <https://doi.org/10.19173/irrodl.v18i1.2623>

- Eke, D. (2023). ChatGPT and the rise of generative AI: Threat to academic integrity? *Journal of Responsible Technology*, 13, 1–4. <https://doi.org/https://doi.org/10.1016/j.jrt.2023.100060>
- Fido, D., & Wallace, L. (2023). The unique role of ChatGPT in closing the awarding gap. *The Interdisciplinary Journal of Student Success*, 1–7.
- Firat, M. (2023). What ChatGPT means for universities: Perceptions of scholars and students. *Journal of Applied Learning & Teaching*, 6(1), 1–7. <https://doi.org/https://doi.org/10.37074/jalt.2023.6.1.22>
- Google News Initiative. (2019). Google Trends: Understanding the data. In *Google*. https://storage.googleapis.com/gweb-news-initiative-training.appspot.com/upload/GO802_NewsInitiativeLessons_Fundamentals-L04-GoogleTrends_1saYVCP.pdf
- Halaweh, M. (2023). ChatGPT in education: Strategies for responsible implementation. *Contemporary Educational Technology*, 15(2), ep421. <https://doi.org/10.30935/cedtech/13036>
- Haleem, A., Javaid, M., & Singh, R. P. (2022). An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges. *BenchCouncil transactions on benchmarks, standards and evaluations*, 2(4), 100089.
- Ina. (2022). *The history of chatbots – From ELIZA to ChatGPT*. <https://onlim.com/en/the-history-of-chatbots/>
- Khalil, M., & Er, E. (2023). Will ChatGPT get you caught? Rethinking of Plagiarism Detection. *ArXiv*, 1–13. <https://doi.org/https://doi.org/10.35542/osf.io/fnh48>
- Kumar, D. (2022). *What is ChatGPT-3? Why it has become so popular?* <https://coingape.com/blog/what-is-chatgpt-3-why-it-has-become-so-popular/>
- Lim, W. M., & Ting, D. H. (2012). E-shopping: An analysis of the technology acceptance model. *Modern Applied Science*, 6(4), 49–62. <https://doi.org/10.5539/mas.v6n4p49>
- McQuail, D., & Windahl, S. (1993). *Communication models for the study of mass communications* (2nd ed). Longman.
- O'Connor, C., & Joffe, H. (2020). Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines. *International Journal of Qualitative Methods*, 19, 1–13. <https://doi.org/10.1177/1609406919899220>
- OpenAI. (2022). *ChatGPT: Optimizing language models for dialogue*. <https://openai.com/blog/chatgpt/>
- OpenAI. (2023). *GPT-4 is OpenAI's most advanced system, producing safer and more useful responses*. <https://openai.com/product/gpt-4>
- Orr, G. (2003). *Diffusion of innovation by Everett Rogers*. [http://www.stanford.edu/class/symsys205/Diffusion of Innovations.htm](http://www.stanford.edu/class/symsys205/Diffusion%20of%20Innovations.htm)
- Rathore, B. (2023). Future of AI & generation alpha : ChatGPT beyond boundaries. *Eduzone*, 12(1), 63–68.
- Rau, G., & Shih, Y. S. (2021). Evaluation of Cohen's kappa and other measures of inter-rater agreement for genre analysis and other nominal data. *Journal of English for Academic Purposes*, 53(July), 101026. <https://doi.org/10.1016/j.jeap.2021.101026>
- Rogers, M. E. (1995). *Diffusion of innovation* (5th ed.). The Free Press.
- Shorten, A., & Smith, J. (2017). Mixed methods research: Expanding the evidence base. *Evidence-Based Nursing*, 20(3), 74–75. <https://doi.org/10.1136/eb-2017-102699>
- Sok, S., & Heng, K. (2023). ChatGPT for education and research: A review of benefits and risks. *SSRN*. <https://doi.org/https://dx.doi.org/10.2139/ssrn.4378735>
- Susnjak, T. (2022). ChatGPT: The end of online exam integrity? *ArXiv*, 1–21. <https://doi.org/https://doi.org/10.48550/arXiv.2212.09292>
- Tinmaz, H. (2011). *Utilization of social networking websites in education: A case of facebook*. Middle East Technical University.

- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, *10*(1). <https://doi.org/10.1186/s40561-023-00237-x>
- Wang, Y., Matz-Costa, C., Miller, J., Carr, D. C., & Kohlbacher, F. (2018). Uses and gratifications sought from mobile phones and loneliness among Japanese midlife and older adults: A mediation analysis. *Innovation in Aging*, *2*(3), 1–13. <https://doi.org/10.1093/geroni/igy027>
- Willems, J. (2023). ChatGPT at universities – The least of our concerns. *SSRN*. <https://doi.org/10.2139/ssrn.4334162>
- Wodzicki, K., Schwammlein, E., & Moskaliuk, J. (2012). “Actually, I wanted to learn”: Study-related knowledge exchange on social networking sites. *Internet and Higher Education*, *15*(1), 9–14. <https://doi.org/10.1016/j.iheduc.2011.05.008>

APPENDIX 1

The Text Generated by Chatgpt

The flipped classroom model has gained popularity in recent years as a way to improve student learning and engagement. This approach involves flipping the traditional model of instruction, where students receive lectures in the classroom and complete assignments at home, by having students watch lectures and complete assignments in class, while using class time for interactive activities and discussions.

One variation of the flipped classroom model is collaborative flipped learning, where students work together in small groups to complete assignments and discuss course material. This approach has been shown to promote active learning and critical thinking, as well as improve social and communication skills.

Studies have found that collaborative flipped learning can lead to increased student achievement and engagement, as well as improved attitudes towards the course and the subject matter. For example, a study by Kim and Park (2013) found that students in a collaborative flipped learning environment had higher test scores and were more engaged in class compared to students in a traditional classroom. Similarly, a study by Huang and Chiu (2014) found that students in a collaborative flipped learning environment had better problem-solving skills and were more confident in their ability to apply what they had learned.

However, not all research on collaborative flipped learning has found positive results. A study by Lee and Lee (2015) found that students in a collaborative flipped learning environment had lower test scores and were less engaged in class compared to students in a traditional classroom. The authors attributed this to the lack of structure and support in the collaborative flipped learning environment, which may have led to confusion and frustration among students.

Despite these mixed findings, collaborative flipped learning has the potential to be an effective approach for improving student learning and engagement. Further research is needed to understand the factors that contribute to its success, as well as to develop strategies for addressing the challenges that may arise in a collaborative flipped learning environment.

Overall, the literature on collaborative flipped learning suggests that this approach has the potential to improve student achievement and engagement, but more research is needed to understand the conditions under which it is most effective. Further investigation into the use of collaborative flipped learning in different subject areas and contexts is necessary to fully understand its potential as a teaching and learning strat