M. AHMED

# THE IMPACT OF CHATGPT ON TASK DIFFICULTY, INTERESTINGNESS, POSITIVE AND NEGATIVE FEELINGS, ACHIEVEMENT, AND RETENTION: A COMPARATIVE STUDY WITH GOOGLE, E-TEXTBOOK, AND NO TOOL

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
ATILIM UNIVERSITY

MOHAMMED AHMED A. ABDULRAZZAQ

A MASTER OF SCIENCE THESIS
IN
THE DEPARTMENT OF COMPUTER ENGINEERING

JUNE 2024

ATILIM UNIVERSITY 2024

# THE IMPACT OF CHATGPT ON TASK DIFFICULTY, INTERESTINGNESS, POSITIVE AND NEGATIVE FEELINGS, ACHIEVEMENT, AND RETENTION: A COMPARATIVE STUDY WITH GOOGLE, E-TEXTBOOK, AND NO TOOL

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
ATILIM UNIVERSITY

BY

MOHAMMED AHMED A. ABDULRAZZAQ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
THE DEPARTMENT OF COMPUTER ENGINEERING

JUNE 2024

Approval of the Graduate School of Natural and Applied Sciences, Atilim University.

_____

Prof. Dr. Ender KESKİNKILIÇ
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of **Master of Science in Computer Engineering, Atılım University**.

_____

Prof. Dr. Gökhan ŞENGÜL
Head of Department

This is to certify that we have read THE IMPACT OF CHATGPT ON TASK DIFFICULTY, INTERESTINGNESS, POSITIVE AND NEGATIVE FEELINGS, ACHIEVEMENT, AND RETENTION: A COMPARATIVE STUDY WITH GOOGLE, E-TEXTBOOK, AND NO TOOL submitted by MOHAMMED AHMED A. ABDULRAZZAQ and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

_____

Assoc. Prof. Dr. Sacip TOKER
Supervisor

**Examining Committee Members:**

Assist. Prof. Dr. Tuncer AKBAY
Information Systems and Technology,
Burdur Mehmet Akif Ersoy University _____

Assoc. Prof. Dr. Sacip TOKER
Information Systems Engineering,
Atilim University _____

Assist. Prof. Dr. Damla TOPALLI
Computer Engineering,
Atilim University _____

**Date:** 24.06.2024

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

MOHAMMED AHMED A. ABDULRAZZAQ

Signature:

# ABSTRACT


## THE IMPACT OF CHATGPT ON TASK DIFFICULTY, INTERESTINGNESS, POSITIVE AND NEGATIVE FEELINGS, ACHIEVEMENT, AND RETENTION: A COMPARATIVE STUDY WITH GOOGLE, E-TEXTBOOK, AND NO TOOL

Mohammed Ahmed

MSc., Department of Computer Engineering

Supervisor: Assoc. Prof. Dr. Sacip Toker

June 2024, 73 pages

This study investigates the impact of ChatGPT 3.0, a sophisticated AI conversational agent, on numerous aspects of student learning outcomes when compared to traditional educational resources such as Google Search and PDF documents. The research focuses on six main areas: quiz scores, task difficulty, task interestingness, positive feelings about the task, negative feelings about the task, and retention. This study featured a group of university students who completed assignments with both ChatGPT and traditional technology. A true-experimental pre- and post-test design was used, with participants randomly assigned to four groups: control, e-textbook, Google, and ChatGPT access. The data reveal that ChatGPT significantly reduces perceived task difficulty across cognitive stages, enhances task interest, and leads to higher quiz scores, especially in the early stages of learning. Furthermore, students reported positive feelings to using ChatGPT, such as feeling less frightened and more supported, while also expressing some negative emotions. These data imply that, while ChatGPT has the potential to significantly improve educational outcomes, its incorporation into educational practices must be carefully addressed to maximize benefits while minimizing potential downsides.

Keywords: ChatGPT, Quiz scores, Task difficulty, Task interestingness, Positive feelings about the task, Negative feelings about the task, Retention, Academic integrity.

# ÖZ

## CHATGPT'NİN GÖREV ZORLUĞU, İLGİ ÇEKİCİLİĞİ, OLUMLU VE OLUMSUZ DUYGULAR, BAŞARI VE ELDE TUTMA ÜZERİNDEKİ ETKİSİ: GOOGLE, E-DERS KİTABI VE ARAÇSIZ KARŞILAŞTIRMALI BİR ÇALIŞMA

Mohammed Ahmed
Yüksek Lisans, Bilgisayar Mühendisliği
Tez Yöneticisi : Doç. Dr. Sacip TOKER

Haziran 2024, 73 sayfa

Bu çalışma, gelişmiş bir yapay zeka konuşma aracısı olan ChatGPT'nin 3.0, Google Arama ve PDF belgeleri gibi geleneksel eğitim kaynaklarıyla karşılaştırıldığında öğrenci öğrenme sonuçlarının çeşitli yönleri üzerindeki etkisini araştırıyor. Araştırma altı ana alana odaklanıyor: sınav puanları, görevin zorluğu, görevin ilgi çekiciliği, göreve ilişkin olumlu duygular, göreve ilişkin olumsuz duygular ve kalıcılık. Bu çalışma, hem ChatGPT hem de geleneksel teknolojiyle ödevlerini tamamlayan bir grup üniversite öğrencisini içeriyordu. Katılımcıların rastgele dört gruba atandığı gerçek deneysel bir ön ve son test tasarımı kullanıldı: kontrol, e-ders kitabı, Google ve ChatGPT erişimi. Veriler, ChatGPT'nin bilişsel aşamalarda algılanan görev zorluğunu önemli ölçüde azalttığını, göreve ilgiyi artırdığını ve özellikle öğrenmenin erken aşamalarında daha yüksek sınav puanlarına yol açtığını ortaya koyuyor. Ayrıca öğrenciler ChatGPT kullanımına ilişkin daha az korkmak ve daha fazla destek görmek gibi olumlu duygusal tepkiler verirken aynı zamanda bazı olumsuz duyguları da ifade ettiklerini bildirdiler. Bu veriler, ChatGPT'nin eğitim sonuçlarını önemli ölçüde iyileştirme potansiyeline sahip olmasına rağmen, potansiyel dezavantajları en aza

indirirken faydaları en üst düzeye çıkarmak için eğitim uygulamalarına dahil edilmesinin dikkatle ele alınması gerektiğini ima ediyor.

Anahtar Kelimeler: ChatGPT, Sınav puanları, Görevin zorluğu, Görevin ilgi çekiciliği, Görevle ilgili olumlu duygular, Görevle ilgili olumsuz duygular, Kalıcılık, Akademik dürüstlük.

*To my father, Ahmed & my mother, Salma*

*I can never thank you enough for your support and belief in me.*

*You are my greatest inspiration and my endless source of strength.*

## ACKNOWLEDGMENTS

I would like to express my sincere thanks to my advisor, Dr. Sacip Toker, for his insightful guidance, patience, and support. His expertise and constructive feedback have been instrumental in the completion of this research.

Beyond the invaluable academic guidance, I would also like to express my deepest gratitude to my family, especially my father, Ahmed, for his unwavering support, and my mother, Salma, for her steadfast belief in me. Their encouragement has been invaluable throughout my academic journey.

I shall also thank my colleagues and friends who provided support and encouragement during my studies. Their assistance and camaraderie have made this journey more manageable and enjoyable.

Furthermore, I thank the members of Defence Jury Committee, Dr. Damla Topalli and Dr. Tuncer Akbay for their valuable feedback and suggestions, which have significantly enhanced the quality of this work.

Finally, I would like to thank all the participants and contributors who directly or indirectly supported this research. Your contributions have been crucial in the successful completion of this project.

**TABLE OF CONTENT**

## LIST OF TABLES

## LIST OF FIGURES

# CHAPTER 1

# INTRODUCTION

## 1.1 Background

The term Artificial Intelligence (AI) in educational settings would have to be a turning point in pedagogy and even the field of technology-based learning, as it generally characterizes an educational environment that incorporates technology. The shift from traditional instructional content to AI-rich platforms is fast evolving in the world of educational technology, ushering in a new era of more personalized, adaptive learning support [1]. AI technologies like ChatGPT are at the center of this evolution, promising to fundamentally change how we learn interactively with rich, relevant, and responsive educational experiences.

In earlier days the key knowledge repositories were textbooks and web-based searches for educational resources. However, textbooks are structured knowledge and that comes with a disadvantage: their inflexibility in accommodating to different learning styles and demands of the students [2]. Contrast that with the Google Search Engine which has an immense pool of knowledge, but often forces learners to look up through tons of information all on their own in a way that requires levels of critical thinking and self-direction that not every student is equipped for [3]. As helpful as traditional methods of learning have been, they indicate a clear need for more flexible and individualized ways of teaching.

New AI technologies, for example ChatGPT are an evolution of the new types of artificial intelligence. where regular static textbooks and unguided research through search engines will often fail to do so AI Solutions like ChatGPT can provide interactive conversational approach for learning capable of tailoring content and feedback in accordance with individual learning pace and comprehension

[4]potentially not just making learning interesting but also personalizing it in manners that have been known to improve understanding and retention rates via tailored instructional interventions.

Although the promise of AI in education is greatly highlighted, recent research comparing the efficacy of this AI-powered approach with more traditional learning resources is still in its infancy. Empirical research suggests that AI-enhanced learning platforms could increase engagement, understanding and retention. One of the points made by Gee evident in his analysis on video games as great learning environments, lead us to remember that game design principles such as engaging players deeply with meaningful problem-solving and rich understanding through active participation, may give an indication about how AI technologies such as ChatGPT could also influence learners. Gee provides an example of how highly interactive platforms that encourage prolonged engagement and the establishment of a connection with content can significantly enhance learning. Considering the application of these ideas in educational AI-based tools, shows some potential for ChatGPT to serve learners more interestingly and with perhaps a better fit to address different learning needs than other static resources[5]. These findings underscore the importance of conducting additional study on how technologies such as ChatGPT affect learners' feelings, perceived task difficulty, perception related to task interestingness, and academic accomplishment when compared to traditional resources.

## 1.2 Research Problem and Justification

Despite the increasing use of Artificial Intelligence (AI) technologies in education, there is a considerable empirical gap in comparing their influence on various student learning outcomes to traditional resources. This study will look at how ChatGPT affects students' feelings about the tool, perceived task difficulty, perception related to task interestingness, academic performance (quiz scores), and retention when compared to traditional resources like Google Search and E-Textbook.

When examined separately, there is an abundance of research on artificial intelligence technologies and traditional educational resources. Comparative studies are less common, even though they are critical for determining the relative benefits and drawbacks of each technique in the same educational situation. Understanding these characteristics is essential for effective learning environments.

Emotional reactions and perceived task difficulty have also been thoroughly studied in terms of their effects on learning outcomes. Affective computing, often known as emotional AI, is becoming increasingly significant in education as it aims to understand and emulate human emotions. These advancements may improve learning processes and outcomes by incorporating emotional intelligence into educational technologies [6]. Emotional intelligence competencies have a major impact on student accomplishment and are critical to the effectiveness of educators [7]. Further study indicates that the appropriate use of AI, like chatbots, may boost student engagement and learning outcomes [8].

It is widely recognized how important student participation and involvement are in achieving academic success. There is yet little known about how AI technologies such as ChatGPT can engage students, which is an area that requires further research. AI systems have been shown in trials to improve teacher-student communication, potentially increasing student participation and involvement [9].

Despite these promising advances, the ethical implications of AI, such as bias, data privacy, and the digital divide, require further investigation. The practical obstacles in implementing AI technology in various educational environments underscore the importance of conducting detailed efficacy evaluations for different student groups [10]. This study will evaluate the effects of ChatGPT on traditional learning resources to provide informative information on how to maximize educational resources and techniques and so improve student learning outcomes.

Objectives of the Study

The primary objective of this study is to assess the role of artificial intelligence (AI) technologies, specifically ChatGPT, in improving educational outcomes when compared with traditional learning media. This study intends to add to the ongoing discussion about the future of educational technology by investigating the effects of AI-powered tools on various aspects of learning and giving practical recommendations for educators, curriculum developers, and technology innovators.

The main goal of this research is to determine the efficacy of ChatGPT in increasing learning outcomes, student engagement, and cognitive development when compared to traditional educational resources such textbooks and web-based searches.

In terms of specific objectives, this study seeks to investigate the impact of ChatGPT on students' feelings during learning activities. This involves assessing how ChatGPT's interactive and responsive nature affects anxiety, motivation, and overall emotional well-being when compared to traditional methods. Another specific goal is to investigate how ChatGPT influences students' perceptions of task difficulty, and whether the quick feedback and individualized support provided by ChatGPT make learning tasks appear less difficult than traditional tools.

The study also aims to assess the influence of ChatGPT on students' interest in learning activities, specifically whether the engaging and interactive character of ChatGPT boosts students' excitement and persistent interest in their studies. Furthermore, the study intends to assess the impact of ChatGPT on academic achievement, with a focus on quiz scores and retention rates, and to compare its effectiveness in enhancing academic outcomes to those of traditional educational methods.

Furthermore, the project will identify and address the practical and ethical issues of implementing AI technologies in education. This includes an examination of topics such as AI reliance, data privacy, academic integrity, and digital divide.

By achieving these goals, the study aims to provide comprehensive insights on the benefits and problems of incorporating AI tools such as ChatGPT into educational procedures. The findings are likely to assist educators, curriculum authors, and technology developers in creating more effective, engaging, and ethical learning interventions that combine the benefits of AI and traditional educational resources.

## 1.3 Research Questions

### 1.3.1 Primary Research Question

The primary research question of this study is: How does the use of ChatGPT compare to traditional learning materials, such as E-Textbook and the Google Search Engine, in affecting students' feelings, perceived task difficulty, perception related to task interestingness, academic performance, and retention?

### 1.3.2 Secondary Research Questions

1. Does ChatGPT influence academic performance, specifically in terms of quiz scores, compared to Google, E-textbook and no tool?

2. What is the impact of ChatGPT on students' perceived task difficulty in comparison with Google, E-textbook and no tool?
3. Does perception related to task interestingness in learning activities differ between students using ChatGPT and those using Google, E-textbook and no tool?
4. Does ChatGPT affect students' Positive feelings about the task during learning activities compared to Google, E-textbook and no tool?
5. Does ChatGPT affect students' Negative feelings about the task during learning activities compared to Google, E-textbook and no tool?
6. Does the ChatGPT affect students' retention rates across different task complexities, such as knowledge acquisition, application, and analysis, compared to Google, E-textbook and no tool?

To present a complete view of its influence on student learning and educational practices, these research questions target both the advantages and disadvantages that would arise with ChatGPT being treated as an educational solution.

### 1.4 Thesis Organization

This thesis is divided into five chapters, described below. Chapter 1 provides an overview of AI in education, followed by the problem description and study objectives. Chapter 2 provides a detailed analysis of AI in education, concentrating on its effects on learning effectiveness, student Interestingness, Feelings and cognitive reactions, as well as the practical and ethical challenges that arise from its implementation. Chapter 3 details the study's research methodology, which included participants, data collection methods, and analysis procedures. Chapter 4 contains thorough data analysis results on ChatGPT's effects on students' feelings, perceived task difficulty, perception related to task interestingness, academic performance, and retention. Chapter 5 assesses the findings and compares them to current literature, focusing on the implications for educational practice and policy. It also addresses the key findings, makes recommendations to educators and institutions, admits the study's limits, and offers areas for future research.

# CHAPTER 2

# LITERATURE REVIEW

## 2.1 Overview of ChatGPT and AI in Education

### 2.1.1 Brief Recap of ChatGPT's Capabilities and General Applications in Education

ChatGPT, developed by OpenAI, uses the GPT architecture to mimic human conversations. It is useful in education for holding conversations, answering follow-up questions, and modifying responses based on user input. [11]. Teachers create practice quiz questions, automate comments on student essays, and even create interactive learning environments that support individualized learning paths using ChatGPT [12].

A major use of ChatGPT in education is its integration as a tutor or content generator in virtual learning environments. According to study, for instance, it helps with language learning by offering practice interactions and real time language usage corrections[13]. Its capacity to produce a lot of instructional material quickly and appropriately for the setting helps to satisfy the different needs of students in different fields [14].

### 2.1.2 Discussion on the Integration of AI Technologies in Educational Settings

Using AI tools like ChatGPT in educational systems is a big step toward more advanced technology learning environments. Integrating AI tools like ChatGPT into education advances technological learning environments by offering scalable and customized learning experiences. However, it raises concerns about educational equity, content accuracy, and the potential reduction of human supervision. [15].

The quality and trustworthiness of content produced by AI have been the main areas of recent progress, which educators continue to be quite concerned about. To be accepted and used successfully, AI systems must be made to link well with learning goals and educational standards. Studies demonstrate ongoing attempts to enhance AI response systems such that they provide knowledge that is both substantively true and pedagogically sound [16].

Moreover, views on artificial intelligence in educational contexts are progressively changing from scepticism to a more sophisticated understanding of AI's potential as a collaborative tool rather than a substitute for human instructors. Support for hybrid models where AI supports human educators while augmenting rather than replacing the human aspects of teaching is growing as educational stakeholders gain a deeper understanding of AI's promise and limitations [17]. Even when AI technologies like ChatGPT show promise for improving teaching methods, integrating them needs to be done carefully taking ethical principles, pedagogical alignment, and the overall effect on the learning process into account. Maintaining these technologies as useful and efficient teaching tools requires continuous development and assessment[14].

## 2.2 Impact of ChatGPT on Student Learning Outcomes in terms of Bloom's taxonomy

### 2.2.1 Quiz Scores and Academic Performance

ChatGPT's effects on quiz scores and general academic performance have been proven ambiguous. According to some research, students who use ChatGPT for revision score higher on tests, suggesting that the tool's ability to offer thorough explanations and prompt feedback can assist students in understanding and remembering material[18]. Conversely, there are worries that students depending too much on AI for answers could develop bad habits that impede learning. Because of this reliance, kids may learn surface knowledge, which is the ability to get answers without really understanding the underlying ideas [19].

Moreover, a thorough analysis comparing the effectiveness of ChatGPT with other digital tools in reinforcing learning material showed that, although ChatGPT can quickly produce a large variety of information, knowledge retention rates change over time. Though the first understanding was great, students often found that more

effective long-term memory came from typical study techniques like note taking and discussions, which ChatGPT does not explicitly encourage [20].

Moreover, instant performance and engagement both greatly increased when teachers added ChatGPT into their teaching methods. Still, traditional reinforcement exercises were required to increase recall, suggesting that ChatGPT should be used as an adjunct to academic learning rather than as a standalone method [21].

### 2.2.2 Task Difficulty and Interest

Researchers have studied the effects of ChatGPT on task difficulty and student interest in detail as educators try to figure out how this AI technology changes the learning environment. Studies show that ChatGPT can alter the complexity and engagement of instructional tasks. For example, using ChatGPT in problem-solving increased student interest and deeper material investigation [22]. Evidence does, however, indicate that the ease with which material may be accessed utilizing ChatGPT may reduce the perceived difficulty of tasks, which could be detrimental to the growth of critical thinking and problem-solving abilities [23].

Textbooks, PDFs, and Google searches are examples of standard teaching materials with several uses. Generally speaking, textbooks are more structured and accurate in their material, but ChatGPT provides a more engaging and participatory learning environment. But whereas textbooks often go through a rigorous peer review process, the accuracy of ChatGPT's information fluctuates and requires users to critically assess it [23]. Comparably to Google search, ChatGPT offers a conversational interface that can result in a more personalized and interesting experience, maybe boosting learning interest and retention more successfully than the usually passive process of looking through search results [23].

## 2.3 Student Perceptions and Feelings Responses

### 2.3.1 Positive Feelings about the tool

In academic contexts, students' impressions of ChatGPT have been generally positive, especially with regard to its quick accessibility and depth of knowledge base. Studies abound demonstrating how much students appreciate ChatGPT's quick feedback and array of functions, which include everything from essay writing to complex

mathematics problem solving. One of the most often mentioned factors in reducing stress and anxiety brought on by schoolwork and deadlines is this quick support [24].

Moreover, surveys and case studies have indicated that using ChatGPT may raise student motivation and satisfaction. For instance, students who were integrated with ChatGPT showed greater excitement for taking part in learning activities, which they attributed to the interactive and interesting characteristics of the AI tool. Particularly evident is this growing interest in subjects like technical or highly specialized courses, where students typically have trouble engaging [13].

### 2.3.2 Negative Feelings about the tool

Using ChatGPT in educational contexts has some challenges and limitations, even with its advantages. Students have voiced doubts on the veracity of the data that ChatGPT offers. Students that depend on the AI for academic support have been confused and mistrustful of it in cases when it delivered false or misleading answers [16].

Another serious concern is depending too much on ChatGPT to finish schoolwork. Both teachers and students worry that making solutions easily available could prevent them from using the materials more deeply and from thinking critically. The effectiveness of learning may be limited and problem-solving skills may be hampered by this dependence [23].

Using ChatGPT in educational settings also raises questions about academic dishonesty. Academic honesty has been called into question by the ease with which students can produce essays and complex answers. Universities and schools are aggressively investigating ways to lower this risk, like creating detection systems that can tell the difference between content produced by AI and student work [15].

## 2.4 Comparative Analysis with Other Digital Tools

### 2.4.1 Detailed Comparison of ChatGPT with Google Search Engines and Electronic Textbooks

Students can use ChatGPT, electronic textbooks, and Google search engines among other tools to support their learning. Every instrument has certain benefits and drawbacks that significantly affect how useful it is in educational settings.

With so much information available from so many sources, Google search engines enable students to look at many viewpoints on any one subject [25]. Though the caliber of the material can differ greatly, students may need to use critical thinking skills to locate reliable sources. When a user lacks the necessary prior information, Google search may result in misunderstandings because it does not offer focused or context specific explanations, unlike ChatGPT [25].

Content from electronic textbooks is organized and verified, and it is often linked to curriculum requirements. They are meant to support a planned learning process and offer a reliable and constant information source [26]. Nevertheless, compared to AI systems like ChatGPT, electronic textbooks are less flexible to meet the needs of particular students [4].

Instantaneous, engaging, and customized answers are what ChatGPT does best, which can increase learning motivation [10]. More individualized learning is offered by it than by static electronic textbooks or a variety of Google search results since it may modify conversations and explanations to the user's comprehension level [27].

### 2.4.2 Assessment of Strengths and Weaknesses Across Different Tools with Respect to User Experience, Content Reliability, and Learning Effectiveness

User Experience: Many students find that ChatGPT's conversational interface is more appealing than the static electronic textbook formats or the disjointed portions from a Google search. Students can stay interested in ChatGPT via its interactive format, which mimics a one-on-one tutoring session [28]. Students who prefer more conventional learning techniques, however, may find it disadvantageous as textbooks often lack the visual components and tactile interaction they often offer [4].

Content Reliability: ChatGPT generates its answers using a combination of licensed data, data produced by human trainers, and publicly available material, which might occasionally lead to errors, unlike electronic textbooks, which are peer-reviewed and edited by professionals [29]. Reliability of Google search results varies widely, hence consumers need to be highly digitally literate to recognize trustworthy sites [30].

Learning Effectiveness: Although comprehensive curriculum coverage made possible by electronic textbooks can be a drawback for methodical study. The ability of ChatGPT to produce a wide range of material and provide explanations of subjects in many ways supports various learning styles, which can enhance retention and understanding. Deep learning may be at risk, nevertheless, if responses are obtained only from ChatGPT without any critical engagement [31]. While self-directed exploration is encouraged by Google search engines, it may overwhelm some students but be beneficial for those who are committed and disciplined [25].

## 2.5 Ethical Considerations and Future Implications

### 2.5.1 Discussion on the Ethical Implications of Using AI Tools like ChatGPT in Educational Contexts

To make sure that AI technologies like ChatGPT benefit every student without unintentionally hurting any of them, serious ethical issues must be resolved. Security and privacy of data is one big issue. Huge amounts of data are necessary for AI systems like ChatGPT to learn and function correctly, hence data breaches and misuse are a given [32]. Confidentiality and integrity of student data must be preserved.

Another moral dilemma is fair access to AI technologies. If socioeconomic classes have different access to these technologies, there is worry that they would exacerbate already existing educational disparities [33]. A major worry is also how AI can impact academic integrity and honesty. It may be challenging to evaluate students' actual comprehension and effort when they can write essays and solve challenging tasks with ChatGPT so easily [34]. New rules and procedures must be developed by educational institutions to guarantee the validity of student work and the use of AI technologies to enhance, not to replace, human intelligence and creativity.

### 2.5.2 Predictions and Recommendations for Future Integration of AI Technologies in Learning Environments

With the ability to enhance learning processes and results, the use of AI technology in education is predicted to pick up speed in the next years. Broad guidelines, however, that tackle the moral, legal, and practical concerns brought forth should guide this integration. Implementing an ethical AI framework that includes standards for data

usage, privacy, accessibility, and academic integrity is advocated by educational institutions [35].

Future recommendations include developing hybrid learning models where AI skills are used to enhance rather than to replace current teaching methods. While maintaining important human components of education, such empathy, judgment, and moral direction, these models could make use of AI's strengths in large data and customized learning [36].

Moreover, to guarantee the responsible and effective use of these tools, the impact of AI in educational contexts should be constantly observed and assessed. Frequently seeking for stakeholder input especially from parents, instructors, and students will help to iteratively improve AI integration strategies [37].

## 2.6 Synthesize How These Findings Relate to The Thesis Objectives and Questions

The substantial literature review conducted in this thesis investigates how ChatGPT and other AI technologies affect education, highlighting important developments in student engagement, academic success, ethical considerations, and student opinions. The actual data collected during the thesis research is combined with these findings in this synthesis to provide a comprehensive picture of how AI technologies are transforming educational experiences.

It has been demonstrated that ChatGPT increases student engagement and lowers perceived difficulty of educational tasks. Consistent with the thesis findings that students using ChatGPT perceived tasks as less difficult than those using traditional tools like e-textbooks and Google search, these studies confirm that ChatGPT significantly reduces barriers to comprehension and task completion, particularly in the early stages of learning [11], [23].

It's more nuanced how ChatGPT affects academic achievement, though. Even although some research reveals an early rise in quiz scores, this benefit usually disappears as tasks become more difficult, suggesting a possible deep learning and memory gap. ChatGPT's can enhance academic performance over the long run beyond surface-level engagement [12].

Equally important are ethical issues when using AI in educational contexts. It raises issues concerning data privacy, equal access, and the potential for academic dishonesty which emphasize both the dangers associated with data handling and the challenge of upholding academic integrity with AI technologies [10], [15].

Moreover, opinions of students toward ChatGPT are nuanced. Even though many individuals appreciate AI's quick help and customized learning paths, there is general worry about the accuracy of the information provided and an over dependence on such technology [24].

# CHAPTER 3

# METHODOLOGY

This chapter describes the research design, including participants, data gathering methods, and analysis procedures. It describes how the study was carried out, including the experimental design for comparing ChatGPT to standard educational tools, the metrics used to assess learning outcomes, and the statistical techniques used to analyze the data.

## 3.1 Participants

Initially, 152 students from an information systems department at a private university in Turkey agreed to take part in the study. However, 24 students were removed because they did not finish all the required tasks, and another five were excluded for failing to follow instructions correctly. Finally, data from 123 students were evaluated.

Participants took two pretests and two self-efficacy questionnaires to evaluate their understanding of data privacy and ChatGPT, as well as their search self-efficacy and experiences. The average score for the data privacy test was 54.81 (SD = 27.10) out of 100, but the average score for the ChatGPT test was 71.29 (SD = 9.32). The mean score on the search self-efficacy exam was 49.52 (SD = 9.22) out of 70.

## 3.2 Research design

We used a true-experimental pre- and post-test. We grouped the participants into four groups based on their search experience, belief in their own search skills, understanding of data protection, and familiarity with ChatGPT. with random assignment, which a variable that could be changed [38]. Initially, four groups were randomly assigned: control, e-textbook, Google, and ChatGPT access. The E-textbook group could only use a PDF version of the textbook, which served as the principal resource in the data privacy course. The Google group used only the Google search

engine during the trial, and students in this group were not allowed to use any other resources. The ChatGPT group only utilized ChatGPT, without access to the e-textbook or Google. The control group did not have access to any of these tools during the experiment.

The experiment was conducted in a laboratory setting under the supervision of the course team, which included the instructor and teaching assistants. We understand that in real-world situations, students frequently have simultaneous access to various resources, such as Google and ChatGPT. Participants were constantly supervised to ensure they followed the study protocol, and deviations from the allocated tools were kept to a minimum. Only five students initially departed from the recommended processes, but they reverted to the proper instruments after being reminded. These five students were excluded from the study at its conclusion. This careful commitment to experimental conditions helped to maintain the internal validity of our findings.

We used Bloom's Revised Taxonomy of Educational[39] to create three data privacy tasks of varied complexity for this study. Task 1 required students to perform an assignment that involved lower order thinking skills such as remembering and comprehending (see Table 3.1). In Task 2, participants completed an activity that aimed to improve medium-order thinking skills, such as applying (see Table 3.2). Task 3 challenged students to identify a problem and provide a solution, which needed higher order thinking skills such as analysis, evaluation, and creativity (see Table 3.3). Participants completed the activities in order: Task 1, Task 2, and Task 3.

Table 3.1 Task 1- remembering and understanding

Please answer the following questions and submit them as a Word file to the Moodle Course site.

1. What is the European Union (EU) General Data Protection Regulation (GDPR)? List the GDPR requirements that must be met by EU organizations when processing personal information, including but not limited to collection, recording, organization, structuring, storing, adaptation, retrieval, use, dissemination, combination, and destruction of personal data.
2. What types of information are protected by the GDPR?
3. What are the main differences and similarities between privacy, security, and ethics?
4. What do cookies mean? And why are they important for privacy?

Table 3.2 Task 2 – applying

Suppose that you work for a company as a privacy and security expert with proficiency on legal and technical aspects of GDRP. The company's management wants to make sure it has policies and procedures in place to protect the privacy of visitors to its website. You have been asked to provide advice to the company about what needs to be included in the policies to ensure that they comply with the EU GDPR. Your advice should address issues, including but not limited to:

1. How much data should the company collect on visitors to its website, and why that data? What information could it discover by tracking visitors' activities on its website?
2. What value would this information provide the company? What are the privacy problems raised by collecting such data?
3. Should the company use cookies? What are the advantages of using cookies for both the company and its website visitors? What privacy issues do they create for the company?
4. Should the company adopt an opt-in or opt-out model of informed consent?

Table 3.3 Task 3 - analyzing, evaluating, and creating

Examine the Optimum Customer Privacy Notice and evaluate whether it complies with the EU GDPR. Base your evaluation on the GDPR requirements, including but not limited to

1. requiring organizations to protect PII (personally identifiable information),
2. requiring organizations to allow users to access all their personal information without charge within one month,
3. requiring organizations to delete personal data (right to be forgotten),
4. requiring organizations to ensure data portability so consumers are not locked into a particular service,
5. requiring organizations to guarantee the right to sue providers for damages or abuse of PII, including class action lawsuits,
6. requiring organizations to have a data protection officer that reports to senior management;
7. requiring explicit consent before collecting data (positive opt-in),
8. requiring organizations to eliminate default opt-in processes,
9. requiring organizations to publish the rationale for data collection and the length of retention
10. requiring organizations to report breaches and hacks within 72 hours,
11. requiring organizations to assume liability for data they share with partners or other firms,
12. requiring organizations to build privacy protections into all new systems (privacy by design),
13. requiring organizations to anonymize data rather than targeting based on intimate, personal profiles;
14. requiring organizations to limit the collection of personal data to only that which is needed to support a task or a transaction and then deleting it shortly thereafter.

After the evaluation, identify gaps in the privacy policy. What are your suggestions for improving the privacy policy (PR)? Please prepare a list of suggestions to improve the PR document.

### 3.3 Group formation

We grouped the participants into four groups based on their search experience, belief in their own search skills, understanding of data protection, and familiarity with ChatGPT. This categorization was accomplished using a two-step cluster analysis. The first cluster of 24 subjects had Minimal composite scores in these areas. The second cluster, which included 32 subjects, had somewhat higher but still poor results. The third cluster, which included 62 participants, had average scores, but the fourth cluster, which included 35 participants, had the highest scores across these factors.

To guarantee balanced representation, participants from each cluster were split evenly among the three experimental groups and one control group. This strategy sought to maintain a consistent mix of individuals, with various levels of experience and expertise in each group based on the clustering criteria. Gender balance was also considered in the distribution procedure. Figure 1 illustrates the distribution procedure and participant clustering characteristics.

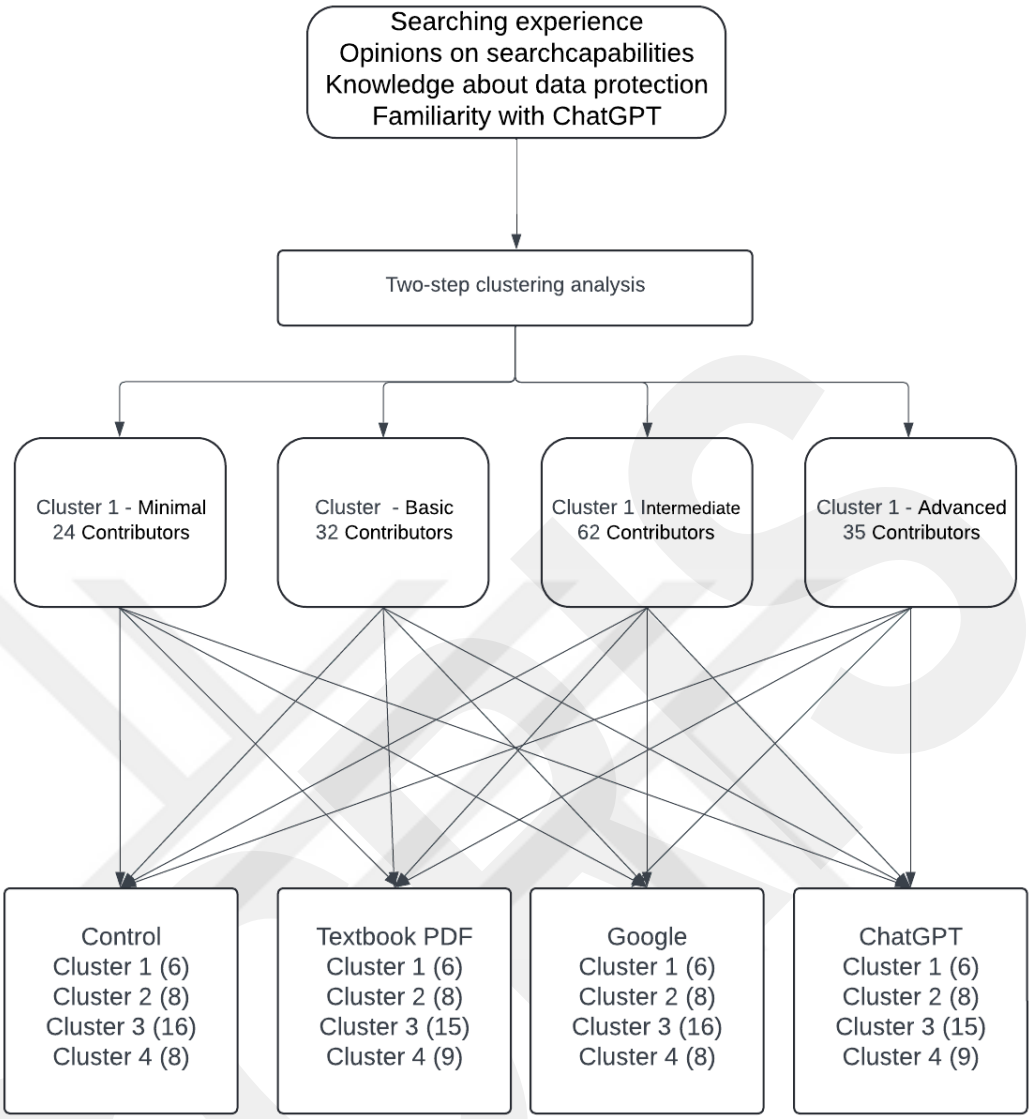


Figure 3.1 The process of random assignment to the study groups

Table 4 shows the distribution of students by group and gender. There were 40 women (32.52%) and 83 men (67.48%) among the participants. The control group had 31 students (25.20%). The treatment groups were 25 students (20.33%) in the e-textbook group, 29 (23.58%) in the Google group, and 38 (30.89%) in the ChatGPT group.

Table 3.4 The frequency distribution of the participants by groups and gender

| Groups | Gender | | | | | |
|---|---|---|---|---|---|---|
| | Female | | Male | | Total | |
| | f | % | f | % | f | % |
| **Control** | 11 | 8.94 | 20 | 16.26 | 31 | 25.20 |
| **E-textbook** | 10 | 8.13 | 15 | 12.20 | 25 | 20.33 |
| **Google** | 9 | 7.32 | 20 | 16.26 | 29 | 23.58 |
| **ChatGPT** | 10 | 8.13 | 28 | 22.76 | 38 | 30.89 |
| **Total** | 40 | 32.52 | 83 | 67.48 | 123 | 100.00 |

### 3.4 Procedure

Prior to giving homework, the teaching team defined plagiarism in the context of the course and provided instances of what constituted plagiarism. The instructions for each activity included the wording shown below.

In our academic environment, plagiarism is defined as the act of using someone else's ideas, work, or expressions without proper acknowledgment, which is considered academic misconduct when submitting assignments. Examples include: using specific phrases or sentences from another source without attribution; copying text directly from books or online materials; closely paraphrasing or translating another work; using facts, statistics, graphs, images, or diagrams without acknowledging their source; and incorporating comments or notes from others, including those from lectures or tutorials with direct quotes. Even copying text from assignment writing services or other external resources without adequate citation constitutes plagiarism. It is critical to employ proper academic referencing to clearly identify the source of any borrowed content in your work.

Each assignment was assigned one lecture/demonstration session (50 minutes) and one lab session (100 minutes) throughout the course, with the lecture/demonstration session coming before the lab session. These sessions included a review of important course content and concepts, as well as demonstrations of how to complete similar tasks. During the lab sessions, students were expected to perform the prescribed assignments.

Following job completion, all entries were uploaded to Turnitin using a learning management system for plagiarism detection. Along with its standard similarity ratings, Turnitin now adds an AI writing detection tool that assesses the percentage of text that is likely written by AI writing software. As a result, each submission was checked for plagiarism using Turnitin, and the data was based on the percentages found. Figure 2 depicts the overall research strategy and procedures used in the study.

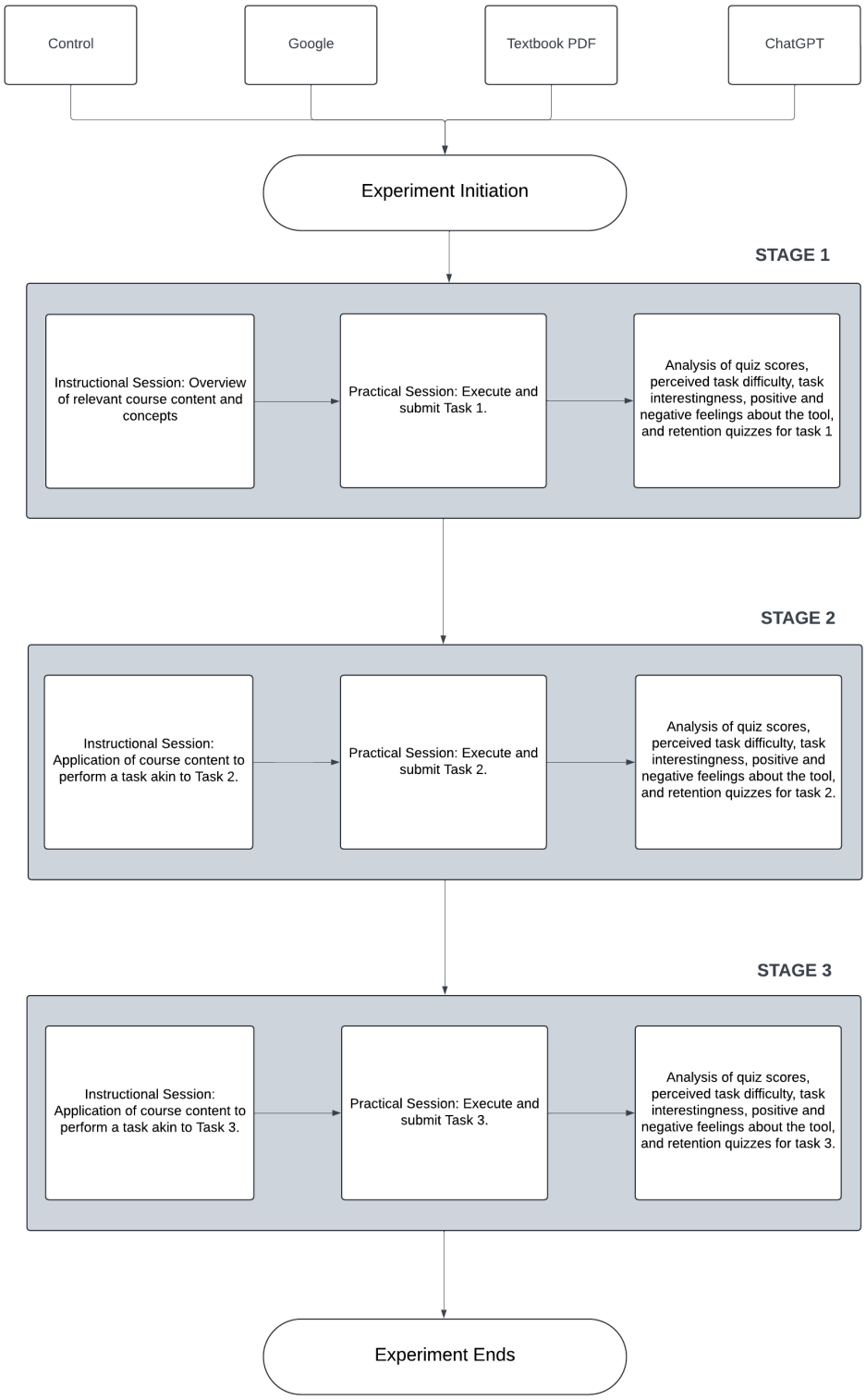


Figure 3.2 Overall research design and data collection

### 3.5 Data Analysis and Measures

To evaluate the influence of ChatGPT on various elements of learning, a variety of measurements and data gathering methods were used. The major indicators were feelings about the tool, perceived task difficulty, Task Interestingness, academic achievement (quiz scores), and retention rates. Each metric is described in full below, along with the data collection process.

#### 3.5.1 Academic Performance (Quiz Scores)

Academic performance was assessed using quizzes built expressly to test knowledge and comprehension of the data privacy issues presented in the tasks. Each quiz contained multiple-choice and short-answer questions designed to check the students' understanding of the content. To confirm the reliability of the academic performance measures, Mauchly's sphericity test was used. The results verified that the data met the sphericity assumptions, hence confirming the validity of the statistical analysis. The Mauchly's sphericity test for quiz scores produced the following results: Mauchly's W = 0.978, $\chi^2(2) = 2.745$, and $p = 0.253$. These findings show that the sphericity assumption was not violated. allowing for a proper interpretation of the repeated measures ANOVA performed on the quiz score data [40]. We ensured the content validity of the assessment tests by focusing on the goals of the tasks and the details of privacy topics. For each task, specific goals were identified in accordance with the complexity. These goals were matched with the privacy topic in detail. The instructor and teaching assistant were ensured for content validity of the quizzes. The reliability of the quizzes was ensured with internal consistency, and KR21 scores for all three assessment was not lower than 0.70. As a result, the assessment tests were both valid and reliable.

#### 3.5.2 Perceived Task Difficulty

Perceived task difficulty was measured using survey items that asked participants to rate the difficulty of every task. Participants rated the perceived difficulty of each task shortly after completion, ensuring that their comments were recorded while the experience was still recent. This method gave an indicator of how difficult participants found each task, depending on the technology they used (e.g., ChatGPT, Google, e-textbook). Mauchly's sphericity test was used to ensure the reliability of the perceived

task difficulty measures. The results verified that the data met the sphericity assumptions, therefore confirming the validity of the statistical analysis. The Mauchly's sphericity test for perceived task difficulty showed the following results: Mauchly's $W = 0.994$, $\chi^2(2) = 0.695$, and $p = 0.707$. The results show that the sphericity assumption was not violated [40].

### 3.5.3 perception related to task interestingness

Perception of task interest were evaluated using survey questions designed to gauge participants' interest and excitement for the assignments. The survey included phrases like "This task is an interesting task" and "I am really curious about the solution of this task." Participants filled out a perception of task interest survey after each task, allowing us to track variations in perception of task interest across activities and learning tools (e.g., ChatGPT, Google, e-textbook). To validate the dependability of the perception of task interest metrics, Mauchly's sphericity test was used. The results verified that the data met the sphericity assumptions, hence confirming the validity of the statistical analysis. The Mauchly's sphericity test for perception of task interest produced the following results: Mauchly's $W = 0.984$; $\chi^2(2) = 2.024$; $p = 0.364$. These findings show that the sphericity assumption was not violated ($p > 0.05$), allowing for a proper interpretation of the repeated measures ANOVA performed on the perception of task interest data.

### 3.5.4 Positive Feelings about the Tool

Positive feelings about the tool were analyzed using a standardized survey designed to capture a variety of positive emotional reactions during the learning exercises. The poll used Likert scale items to examine feelings of confidence, enthusiasm, and pleasure with the tool. Participants assessed each remark on a scale of "Disagree strongly" to "Agree strongly," providing an objective estimate of their positive feelings about the tool. The survey was given right after each task to ensure that participants' replies reflected their immediate positive emotional reactions to the learning activities. The reliability of the positive feelings measures was validated using Mauchly's sphericity test. The results showed that the sphericity assumption was not met: Mauchly's $W = 0.830$, $\chi^2(2) = 22.986$, $p < 0.001$. Because the p-value was less than 0.05, the assumption of sphericity was not met, requiring modifications to the degrees

of freedom for the repeated measures ANOVA. To account for this violation, the Greenhouse-Geisser ($\varepsilon = 0.854$) and Huynh-Feldt ($\varepsilon = 0.886$) corrections were used. This statistical adjustment ensures the reliability of our findings on the positive emotional impact of using various learning technologies on students.

### 3.5.5 Negative Feelings about the Tool

Negative feelings about the tool were measured using a standardized survey designed to capture a range of negative emotional reactions during the learning activities. The poll used Likert scale items to examine feelings like anxiety, worry, and unhappiness with the technology. Participants assessed each remark on a scale of "disagree strongly" to "agree strongly," providing an objective indication of their negative emotional reactions. The survey was given just after each task to ensure that participants' replies reflected their immediate negative emotional reactions to the learning activities. The reliability of the negative feelings measures was validated using Mauchly's sphericity test. The results confirmed that the data met the assumptions of sphericity, as demonstrated by the following results: Mauchly's $W = 0.969$; $\chi^2(2) = 3.919$; $p = 0.141$. These findings show that the sphericity assumption was not broken ($p > 0.05$), allowing for a proper interpretation of the repeated measures ANOVA performed on the negative sentiments data.

### 3.5.6 Retention quizzes

A follow-up questionnaire was administered two weeks after the last task to determine retention rates. The follow-up quiz was created to test the same concepts as the initial quizzes, allowing us to determine long-term retention of the subject. The follow-up quiz scores were compared to the original quiz scores to calculate each participant's retention rates. This comparison provides a quantitative assessment of how well students retained the material they received, based on the tools they used (e.g., ChatGPT, Google, e-textbook).

When we ran the test to compare between Retention and Quiz scores we found that for Task 1 it gave Sphericity Assumed gave (F-value = 5.68, p-value = 0.001) and Wilks' Lambda gave (F-value = 5.68, p-value = 0.001) which indicates that our results are statistically significant but in Task 2 Sphericity Assumed (F-value = 0.47, p-value = 0.702 ) , And Wilks' Lambda gave (F-value = 0.47, p-value = 0.702) and Task 3

Sphericity Assumed (F-value = 0.49, p-value = 0.689) , And Wilks' Lambda (F-value = 0.49, p-value = 0.689) that means that our results are not statistically significant for Task 3 and Task 2 since it exceeds the generally used 0.05 threshold.

# CHAPTER 4

# RESULTS

The results chapter offers detailed data collecting analyses together with the study's conclusions. It comprises statistical findings on how ChatGPT affects students' emotional reactions, perceived task difficulty, perception related to task interestingness, academic performance, retention, and emotional response This chapter also provides tables and figures to help explain the key findings.

## 4.1 Quiz Scores

This study examines the impact of various informational and technological tools, including Google Search Engine, ChatGPT, and PDF documents, on students' quiz scores to understand their effectiveness in enhancing academic performance.to do that Mauchly's Test of Sphericity was used to evaluate the assumption of equal variances in phase differences for the QuizScores metric. The test yielded a result of (Mauchly's W =0.978, $\chi^2(2) = 2.745$, p-= 0.253). This p-value is greater than the standard alpha criterion of 0.05, indicating that the assumption of sphericity for the quiz score data has not been violated.

Table 4.1 Within-Subjects Effects Analysis for Quiz Scores

| **Tests of Within-Subjects Effects** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Measure: | Quiz Scores | | | | | | | | |
| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared | Noncent. Parameter | Observed Power[a] |
| Task | Sphericity Assumed | 56162.49 | 2 | 28081.24 | 145.46 | 0.00 | 0.54 | 290.91 | 1.00 |
| | Greenhouse-Geisser | 56162.49 | 1.96 | 28701.07 | 145.46 | 0.00 | 0.54 | 284.63 | 1.00 |
| | Huynh-Feldt | 56162.49 | 2.00 | 28081.24 | 145.46 | 0.00 | 0.54 | 290.91 | 1.00 |
| | Lower-bound | 56162.49 | 1.00 | 56162.49 | 145.46 | 0.00 | 0.54 | 145.46 | 1.00 |
| Task * Groups | Sphericity Assumed | 9940.95 | 6 | 1656.83 | 8.58 | 0.00 | 0.17 | 51.49 | 1.00 |
| | Greenhouse-Geisser | 9940.95 | 5.87 | 1693.39 | 8.58 | 0.00 | 0.17 | 50.38 | 1.00 |

Table 4.1 (Continued)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Huynh-Feldt | 9940.95 | 6.00 | 1656.83 | 8.58 | 0.00 | 0.17 | 51.49 | 1.00 |
| | Lower-bound | 9940.95 | 3.00 | 3313.65 | 8.58 | 0.00 | 0.17 | 25.75 | 0.993 |
| Error(Task) | Sphericity Assumed | 47877.64 | 248 | 193.06 | | | | | |
| | Greenhouse-Geisser | 47877.64 | 242.64 | 197.32 | | | | | |
| | Huynh-Feldt | 47877.64 | 248.00 | 193.06 | | | | | |
| | Lower-bound | 47877.64 | 124.00 | 386.11 | | | | | |

In our study, we looked at how quiz results differed among stages. The examination of within-subjects effects (Table 4.1) indicated a very significant main effect of stage, with an F-value of ($F(2, 24) = 145.45$) and a high probability level ($p < 0.001$). The stage had a significant effect size ($\eta^2$) of 0.54, accounting for almost 54% of the variance in quiz scores. These results clearly indicate that the stage of the task has a significant impact on quiz scores.

Table 4.2 Descriptive Statistics for Quiz Scores

| **Descriptive Statistics** | | | | |
|---|---|---|---|---|
| Study Groups | | Mean | Std. Deviation | N |
| QUIZ in Task 1 – Knowing and understanding | My memory | 60.22 | 17.63 | 31 |
| | E-textbook | 66.47 | 20.69 | 27 |
| | Google | 78.99 | 17.66 | 32 |
| | ChatGPT | 82.60 | 17.76 | 38 |
| | Total | 72.87 | 20.36 | 128 |
| QUIZ in Task 2 – Applying | My memory | 47.74 | 19.78 | 31 |
| | E-textbook | 61.85 | 22.36 | 27 |
| | Google | 66.25 | 16.21 | 32 |
| | ChatGPT | 67.89 | 20.02 | 38 |
| | Total | 61.32 | 20.97 | 128 |
| QUIZ in Task 3 - Synthesizing, Evaluating, and Creating | My memory | 46.00 | 14.70 | 31 |
| | E-textbook | 41.70 | 14.64 | 27 |
| | Google | 41.22 | 15.04 | 32 |
| | ChatGPT | 41.15 | 14.77 | 38 |
| | Total | 42.46 | 14.76 | 128 |

The descriptive statistics for quiz scores across three different stages (Table 4.2) indicate how various tools may have impacted student performance:

In QUIZ in Task 1 – Knowing and understanding, students using 'ChatGPT' scored the highest on average (M = 82.60), followed by those using 'Google' (M = 78.99). 'E-

textbook' users had a lower average score (M = 66.46), while the control group scored the lowest (M = 60.21). The total average score for this stage was 72.87.

During QUIZ in Task 2 – Applying, the average scores declined across all tools. 'ChatGPT' users maintained the highest scores (M = 67.89), closely followed by 'Google' users (M = 66.25). 'E-textbook' users showed some improvement from the previous stage (M = 61.85), whereas control group continued to score lower (M = 47.74). The overall average for this stage was 61.33.

By QUIZ in Task 3 - Synthesizing, Evaluating, and Creating, there was a notable drop in average scores for all tools. Control group had an average score of (M = 46.01). 'E-textbook', 'Google', and 'ChatGPT' users had very similar average scores (M = 41.70, M = 41.22, and M = 41.16, respectively). The total average for this stage decreased to 42.46.

These statistics suggest that the performance as measured by quiz scores was higher in the earlier stage (Task 1) when more interactive tools like 'Google' and 'ChatGPT' were used, compared to the 'E-textbook' and control group. However, by Task 3, the average quiz scores converged across all groups, indicating that the type of tool had less of an impact on scores as the tasks progressed in complexity or as the novelty of the tools wore off.

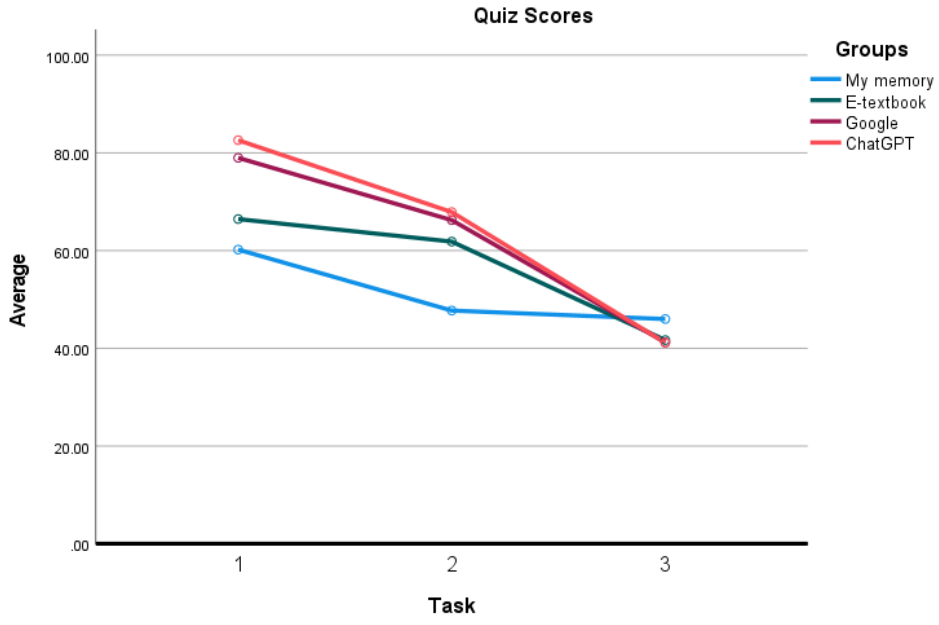


Figure 4.1 Quiz Score Trends by Group and Task

The graph (Figure 4.1) shows the Average means of quiz results over three phases for various study methods: control group, 'E-textbook,' 'Google,' and 'ChatGPT.' Initially, at Stage 1, students utilizing 'ChatGPT' earned the greatest marks, followed by those using 'Google.' The scores for control group and 'E-textbook' were lower. As students advanced through Stages 2 and 3, quiz results for all approaches decreased. 'ChatGPT' and 'Google' users declined, although these two methods maintained the highest scoring throughout. Students in the 'Control group' and those who used the 'E-textbook' also had lower scores, with 'E-textbook' declining more dramatically by Task 2. The converging lines toward the end stage suggest a reduction in the variance in scores between the various research techniques, with 'ChatGPT' maintaining a slight advantage.

### 4.2 Task Difficulty

In our study, we used four distinct tools to assess students' perceptions of how difficult activities were: Control group, E-textbooks, Google, and ChatGPT. We examined three phases of tasks (stages 2, 3, and 4) to see whether the tool they used made the tasks appear simpler or harder, and whether the difficulty level varied depending on the stage of the task they were in.

We ran Mauchly's Test of Sphericity to make sure that the variance in perceived task difficulty across different stages was even. The test showed a result of (Mauchly's W =0.994, $\chi^2(2) = 0.695$ , p- = 0.707). Because this p-value is above the typical threshold of 0.05, it indicates that there wasn't a significant difference in the variances, meaning our assumption of evenness across stages holds up. This means we can rely on the F-ratios from our repeated measures ANOVA without needing to adjust them for any discrepancies [40].

Table 4.3 Within-Subjects Effects Analysis for Task Difficulty

| **Tests of Within-Subjects Effects** | | | | | | | |
|---|---|---|---|---|---|---|---|
| Measure: | Task Difficulty | | | | | | |
| Source | | Type III Sum of Squares | Df | Mean Square | F | Sig. | Partial Eta Squared |
| Task | Sphericity Assumed | 101.07 | 2 | 50.53 | 28.82 | 0.00 | 0.18 |
| | Greenhouse-Geisser | 101.07 | 1.98 | 50.82 | 28.82 | 0.00 | 0.18 |
| | Huynh-Feldt | 101.07 | 2.00 | 50.53 | 28.82 | 0.00 | 0.18 |
| | Lower-bound | 101.07 | 1.00 | 101.07 | 28.82 | 0.00 | 0.18 |
| Task* Study Groups | Sphericity Assumed | 20.49 | 6 | 3.41 | 1.94 | 0.07 | 0.04 |
| | Greenhouse-Geisser | 20.49 | 5.96 | 3.43 | 1.94 | 0.07 | 0.04 |
| | Huynh-Feldt | 20.49 | 6.00 | 3.41 | 1.94 | 0.07 | 0.04 |
| | Lower-bound | 20.49 | 3.00 | 6.83 | 1.94 | 0.12 | 0.04 |
| Error(Stage) | Sphericity Assumed | 434.86 | 248 | 1.75 | | | |

Table 4.4 ( Continued)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Greenhouse-Geisser | 434.86 | 246.61 | 1.76 | | | |
| | Huynh-Feldt | 434.86 | 248.00 | 1.75 | | | |
| | Lower-bound | 434.86 | 124.00 | 3.50 | | | |

In our study, when we looked at how task difficulty changed across different stages (Table 4.3), we found a significant difference. Specifically, the tests showed a strong effect ($F(2, 24) = 28.82$) with a very low chance of it being random ($p < 0.001$) and a large impact size ($\eta 2$ of 0.189) [40]. This tells us that the stage of the task really does affect how difficult it seems.

Table 4.5 Descriptive Statistics for Task Difficulty

| **Descriptive Statistics** | | | | |
|---|---|---|---|---|
| Study Groups | | Mean | Std. Deviation | N |
| TASK_DIFFICULTY in Task 1 – Knowing and understanding | My memory | 6.09 | 1.88 | 31 |
| | E-textbook | 5.40 | 1.55 | 27 |
| | Google | 5.12 | 1.53 | 32 |
| | ChatGPT | 4.68 | 1.74 | 38 |
| | Total | 5.28 | 1.75 | 128 |
| TASK_DIFFICULTY in Task 2 – Applying | My memory | 6.32 | 1.85 | 31 |
| | E-textbook | 5.77 | 1.55 | 27 |
| | Google | 6.06 | 1.83 | 32 |
| | ChatGPT | 4.63 | 1.73 | 38 |

Table 4.6 (Continued)

| | | | | |
|---|---|---|---|---|
| | Total | 5.64 | 1.86 | 128 |
| TASK_DIFFICULTY in Task 3 - Synthesizing, Evaluating, and Creating | My memory | 6.74 | 2.04 | 31 |
| | E-textbook | 7.03 | 1.87 | 27 |
| | Google | 6.96 | 1.89 | 32 |
| | ChatGPT | 5.50 | 1.72 | 38 |
| | Total | 6.49 | 1.96 | 128 |

In our study, we discovered that control group perceived the tasks to be the most challenging, as seen by the highest average difficulty scores across all stages (Table 4.4). On the other hand, using ChatGPT seems to make activities feel the easiest, implying that it is a useful tool for students. When we compared all the methods, the most difficult to easiest was control group, followed by E-textbooks, Google, and lastly ChatGPT as the most useful. Looking closer within each group, e-textbooks appeared to grow less useful as activities became more sophisticated. The effectiveness of memory and Google fluctuated, indicating that their usefulness may vary depending on the task stage. Overall, difficulty scores increased across stages for all groups, showing that activities became more challenging as students advanced. However, this rise was the smallest for those who used ChatGPT, showing its potential as an educational support tool.

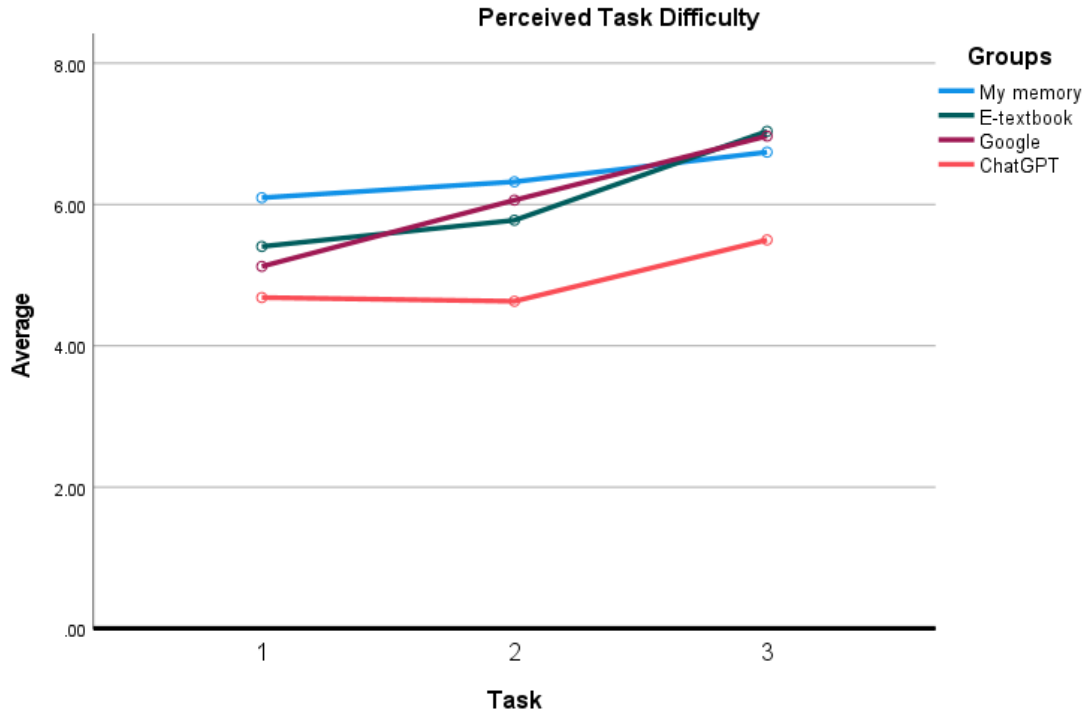


Figure 4.2 Perceived Task Difficulty Trends by Groups and Task

Looking at how students rated the difficulty of tasks in (Figure 4.2), it's evident that control group made problems appear the most challenging. This feeling of hardship increased with each stage. Using an E-textbook seemed to help initially, but toward the end, the activities felt the same, if not more difficult. Google was a bit of a wild card, starting off easy but then becoming more difficult, eventually surpassing the E-textbook.

Conversely, ChatGPT was regarded as the most innovative and popular tool available. It began by making activities appear simple and kept them feeling more controllable than others, despite a modest increase in difficulty later.

In the big scheme of things, all the instruments indicated that tasks were more difficult as students progressed through the phases. But the tale wasn't the same for everyone. The control group had a steep climb in difficulty, whereas the ChatGPT users had a much softer slope, implying that ChatGPT may be the steady hand students require for more difficult tasks.

When we added up all the ratings, ChatGPT came out on top as the most useful tool for reducing perceived difficulty. The control group, however, presented a significant

challenge, as students found the tasks more difficult when required to complete them independently without assistance.

### 4.3 Task interestingness

This study investigates how various informational and technological tools, such as the Google Search Engine, ChatGPT, and PDF documents, influence students' interest in tasks.

We used Mauchly's Test of Sphericity to evaluate the assumption of equal variances in variations between phases for reported task interestingness. The test produced a result of (Mauchly's W =.984, $\chi^2(2) = 2.024$, p- = 0.364). This p-value exceeds the standard alpha level of 0.05, indicating that the assumption of sphericity has not been violated.

Table 4.7 Within-Subjects Effects Analysis for Task interestingness

| **Tests of Within-Subjects Effects** | | | | | | | |
|---|---|---|---|---|---|---|---|
| Measure: | Task Interestings | | | | | | |
| Source | | Type III Sum of Squares | Df | Mean Square | F | Sig. | Partial Eta Squared |
| Task | Sphericity Assumed | 46.38 | 2 | 23.19 | 8.29 | 0.000 | 0.063 |
| | Greenhouse-Geisser | 46.38 | 1.96 | 23.57 | 8.29 | 0.00 | 0.06 |
| | Huynh-Feldt | 46.38 | 2.00 | 23.19 | 8.29 | 0.00 | 0.06 |

Table 4.8 (Continued)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Lower-bound | 46.38 | 1.00 | 46.38 | 8.29 | 0.00 | 0.06 |
| Task * Study Groups | Sphericity Assumed | 9.16 | 6 | 1.52 | 0.54 | 0.77 | 0.01 |
| | Greenhouse-Geisser | 9.16 | 5.90 | 1.55 | 0.54 | 0.77 | 0.01 |
| | Huynh-Feldt | 9.16 | 6.00 | 1.52 | 0.54 | 0.77 | 0.01 |
| | Lower-bound | 9.16 | 3.00 | 3.05 | 0.54 | 0.65 | 0.01 |
| Error(Stages) | Sphericity Assumed | 693.03 | 248 | 2.79 | | | |
| | Greenhouse-Geisser | 693.03 | 244.01 | 2.84 | | | |
| | Huynh-Feldt | 693.03 | 248.00 | 2.79 | | | |
| | Lower-bound | 693.03 | 124.0 | 5.58 | | | |

In our investigation into the variation of task interest throughout different stages, our analysis revealed a significant effect (Table 4.5). The tests of within-subjects effects demonstrated a notable influence ($F(2, 24) = 8.29$) with an extremely low probability of this finding being due to chance ($p < 0.001$). Additionally, the size of the effect was moderately substantial with a partial eta squared ($\eta^2$) of 0.06. These results suggest that the stage of the task has a significant impact on students' perception related to task interestingness.

Table 4.9 Descriptive Statistics for Task interestingness

| **Descriptive Statistics** | | | | |
|---|---|---|---|---|
| Study Groups | | Mean | Std. Deviation | N |
| TASK_INTERESTINGNESS in Task 1 – Knowing and understanding | My memory | 9.93 | 2.81 | 31 |
| | E-textbook | 9.81 | 2.07 | 27 |
| | Google | 10.50 | 1.96 | 32 |
| | ChatGPT | 10.55 | 2.43 | 38 |
| | Total | 10.23 | 2.35 | 128 |
| TASK_INTERESTINGNESS in Task 2 – Applying | My memory | 9.96 | 2.96 | 31 |
| | E-textbook | 9.55 | 2.25 | 27 |
| | Google | 10.56 | 1.81 | 32 |
| | ChatGPT | 9.81 | 2.06 | 38 |
| | Total | 9.98 | 2.30 | 128 |
| TASK_INTERESTINGNESS in Task 3 - Synthesizing, Evaluating, and Creating | My memory | 9.38 | 3.17 | 31 |
| | E-textbook | 8.92 | 2.52 | 27 |
| | Google | 9.59 | 2.74 | 32 |
| | ChatGPT | 9.57 | 2.50 | 38 |
| | Total | 9.39 | 2.72 | 128 |

In this study, we looked at how different tools affected student interest at various cognitive stages (Table 4.6). We discovered that the tools students utilized influenced their perception related to task interestingness. Notably, 'Google' and 'ChatGPT'

received higher initial perception related to task interestingness, indicating that these tools could improve task Interestingness.

In Task 1, 'ChatGPT' had the highest average interest (M = 10.55), closely followed by 'Google' (M = 10.50). The control group generated significant curiosity (M = 9.94), whereas 'E-textbook' had the lowest average (M = 9.81).

In Task 2, 'Google' witnessed a modest increase in interest (M = 10.56), while 'ChatGPT' experienced a tiny decrease (M = 9.82). The control group maintained a very consistent perception related to task interestingness (M = 9.97), while 'E-textbook' remained the lowest average (M = 9.56).

By Task 3, there was a general fall in interest scores across all tools, with The control group having the lowest average interest (M = 9.39) and 'Google' and 'ChatGPT' having roughly comparable perception related to task interestingness (M = 9.59 and 9.58, respectively). Interest in 'e-textbooks' fell even further (M = 8.93).

Overall, while 'Google' and 'ChatGPT' were seen as the most intriguing tools, there was a considerable drop in interest by Task 3, which may reflect the rising difficulty of activities requiring higher-order thinking. According to the findings, choosing the right tool becomes increasingly important as students' progress through more complicated stages of learning.

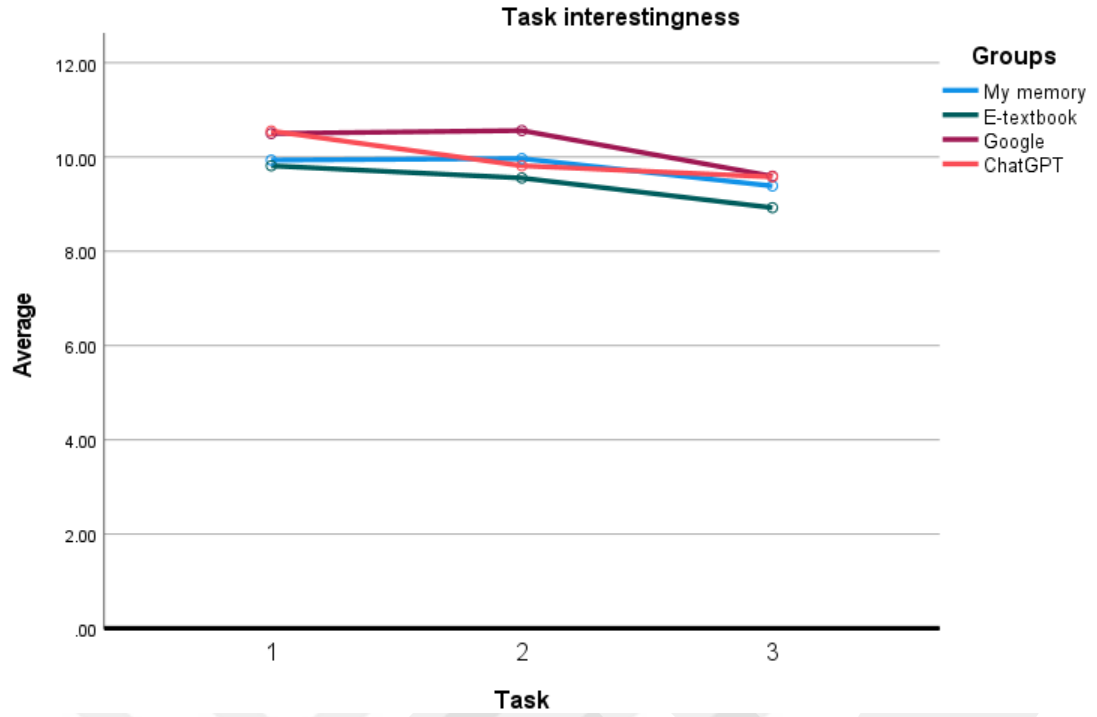


Figure 4.3 Task interestingness Trends by Group and Task

An analysis of task interest over three stages (Figure 4.3) revealed a drop in perception related to task interestingness across all groups from Stage 1 to Task 2. Initially, students who used 'Google' and 'ChatGPT' reported the highest levels of interest, with averages above 10, whereas the control and E-textbooks groups reported slightly lower levels of interest. Despite a little fall, the 'Google' group retained the highest level of interest in Task 2. In contrast, the 'E-textbook' group experienced the biggest decline in interest, yielding the lowest mean score. The 'ChatGPT' group suffered a reduction in interest like the control group, and both groups converged at the end stage. These statistics show that interest in assignments varies depending on the medium used, with digital technology initially engaging students more than traditional approaches. However, the general decline in task interest over phases highlights the need to investigate techniques to keep students involved over time.

## 4.4 Positive Feelings about the tool

This study explores the effects of various informational and technological tools, such as Google Search Engine, ChatGPT, and PDF documents, on students' positive perceptions and feelings towards these resources. We used Mauchly's Test of

Sphericity to determine that the variation in positive feelings across stages was equal. The test yielded a result of (Mauchly's W = 0.830, $\chi^2(2) = 22.986$, $p < 0.001$). Because this p-value is less than the standard threshold of 0.05, it suggests a substantial difference in variances, implying that our assumption of evenness between stages is not valid. Because of the violation of the sphericity assumption, we must rely on modified F-ratios from our repeated measures ANOVA. The Greenhouse-Geisser correction ($\varepsilon = 0.854$) and the Huynh-Feldt correction ($\varepsilon = 0.886$) will be used to offer more accurate F-values and ensure the validity of our results.

Table 4.10 Within-Subjects Effects Analysis for Positive Feelings about the tool

| **Tests of Within-Subjects Effects** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Measure: | Positive Feelings about the tool | | | | | | | | |
| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared | Noncent. Parameter | Observed Power[a] |
| Task | Sphericity Assumed | 55.81 | 2 | 27.90 | 4.00 | 0.01 | 0.03 | 8.01 | 0.71 |
| | Greenhouse-Geisser | 55.81 | 1.70 | 32.66 | 4.00 | 0.02 | 0.03 | 6.84 | 0.66 |
| | Huynh-Feldt | 55.81 | 1.77 | 31.49 | 4.00 | 0.02 | 0.03 | 7.09 | 0.67 |

Table 4.11 (Continued)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Lower-bound | 55.81 | 1.00 | 55.81 | 4.00 | 0.04 | 0.03 | 4.00 | 0.51 |
| Task * Study Groups | Sphericity Assumed | 42.62 | 6 | 7.10 | 1.02 | 0.41 | 0.02 | 6.11 | 0.40 |
| | Greenhouse-Geisser | 42.62 | 5.12 | 8.31 | 1.02 | 0.40 | 0.02 | 5.22 | 0.36 |
| | Huynh-Feldt | 42.62 | 5.31 | 8.01 | 1.02 | 0.40 | 0.02 | 5.42 | 0.37 |
| | Lower-bound | 42.62 | 3.00 | 14.20 | 1.02 | 0.38 | 0.02 | 3.05 | 0.27 |
| Error(stage) | Sphericity Assumed | 1727.85 | 248 | 6.96 | | | | | |
| | Greenhouse-Geisser | 1727.85 | 211.88 | 8.15 | | | | | |
| | Huynh-Feldt | 1727.85 | 219.74 | 7.86 | | | | | |
| | Lower-bound | 1727.85 | 124.00 | 13.93 | | | | | |

The Tests of Within-Subjects Effects (Table 4.7) assesses the effect of stage and the interaction of stage and Study Groups on happy sentiments. The main effect of stage is significant ($F(2, 24) = 4.00$, $p = 0.01$), while the partial eta squared is 0.03, indicating a small effect size. Several corrections (Greenhouse-Geisser, Huynh-Feldt, and Lower-bound) produce substantial findings, demonstrating the robustness of this finding.

Table 4.12 Descriptive Statistics for Positive Feelings about the tool

| **Descriptive Statistics** | | | | |
|---|---|---|---|---|
| Study Groups | | Mean | Std. Deviation | N |
| POSITIVE_FEELING_ABOUT_TOOL in Task 1 – Knowing and understanding | My memory | 16.45 | 4.63 | 31 |
| | E-textbook | 16.11 | 3.12 | 27 |
| | Google | 18.40 | 3.14 | 32 |
| | ChatGPT | 20.50 | 2.94 | 38 |
| | Total | 18.07 | 3.90 | 128 |
| POSITIVE_FEELING_ABOUT_TOOL in Task 2 – Applying | My memory | 17.09 | 4.56 | 31 |
| | E-textbook | 15.55 | 3.51 | 27 |
| | Google | 18.53 | 3.31 | 32 |
| | ChatGPT | 19.71 | 3.00 | 38 |
| | Total | 17.90 | 3.89 | 128 |
| POSITIVE_FEELING_ABOUT_TOOL in Task 3 - Synthesizing, Evaluating, and Creating | My memory | 16.38 | 5.36 | 31 |
| | E-textbook | 14.48 | 4.40 | 27 |
| | Google | 18.25 | 3.95 | 32 |
| | ChatGPT | 18.84 | 3.62 | 38 |

Table 4.13 (Continued)

| | | | | |
|---|---|---|---|---|
| | Total | 17.17 | 4.60 | 128 |

The descriptive statistics for positive feelings about tools (Table 4.8) throughout three stages reveal unique trends among the groups. In Task 1, ChatGPT users had the highest average pleasant feelings (20.50), followed by Google (18.40), My memory (16.45), and E-textbook (16.11). This pattern continues in Task 2, with ChatGPT (19.71) and Google (18.53) having higher means than My memory (17.09) and E-textbook (15.55). In Task 3, ChatGPT (18.84) and Google (18.25) remain ahead, with My memory (16.38) and E-textbook (14.48) trailing. Overall, ChatGPT constantly displays the most pleasant feelings across all stages, with Google also performing admirably. In contrast, E-textbooks have the lowest mean scores, notably in Task 3, showing a less positive perception with time. The standard deviations indicate reasonably similar feelings within groups, with slightly larger variability in the My memory group. These findings show that digital tools, particularly ChatGPT, are perceived more positively than Google and e-textbooks.

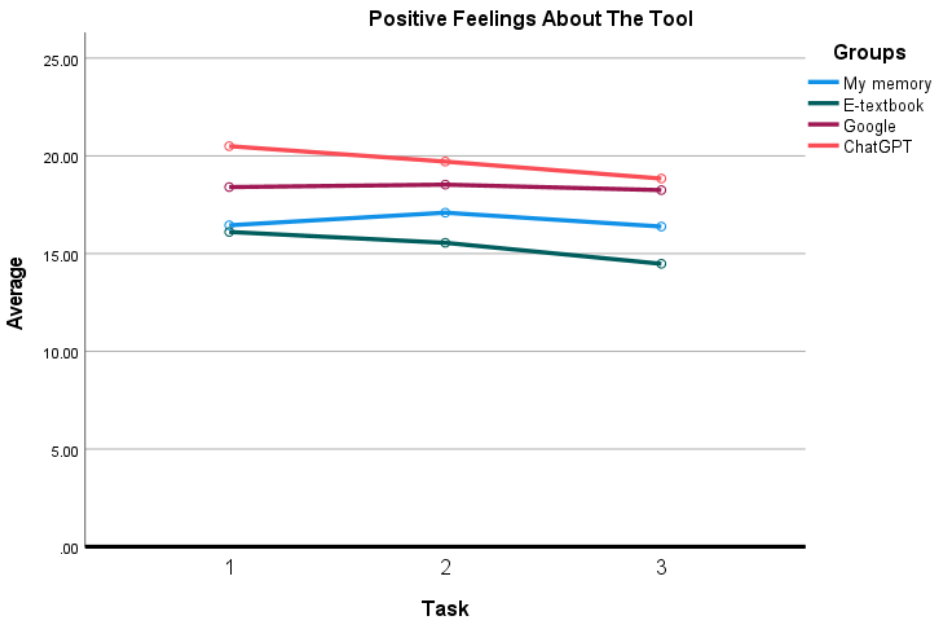


Figure 4.4 Positive Feelings about the tool Trends by Group and Task

The graph of the Average means of positive feelings over three stages reveals distinct trends for each tool (Figure 4.4). ChatGPT constantly receives the most pleasant feelings, beginning above 20 and decreasing slightly between stages. Google also exhibits high levels of positivity, beginning somewhat lower than ChatGPT and gradually decreasing. My memory and E-textbook begin with similar happy sensations, around 16, but My memory improves slightly in Task 2 before dropping in Task 3, but E-textbook declines steadily throughout. Overall, the data indicates that digital tools such as ChatGPT and Google are seen more positively than traditional approaches, with ChatGPT continuously receiving the top rating. These patterns demonstrate a consistent positive perception of AI-based solutions in educational contexts when compared to more traditional study aids.

## 4.5 Negative Feelings about the tool

This study investigates how informational and technological tools like Google Search Engine, ChatGPT, and PDF documents affect students' negative perceptions as they engage with tasks spanning Bloom's Taxonomy.

Mauchly's Test of Sphericity was used to assess the assumption of equal variances in phase differences for the Negative Feelings metric. The test produced a result of Mauchly's W =0.969, ($\chi$^2(2) = 3.919, p = 0.141).

Because the p-value is bigger than the standard alpha level of 0.05, the result is not statistically significant, indicating that the sphericity assumption was not broken for the Negative Feelings variable throughout the stages of your investigation.

Table 4.14 Within-Subjects Effects Analysis for Negative Feelings about the tool

| **Tests of Within-Subjects Effects** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Measure: | Negative Feelings about the tool | | | | | | | | |
| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared | Noncent. Parameter | Observed Power[a] |
| Task | Sphericity Assumed | 62.53 | 2 | 31.26 | 3.71 | 0.02 | 0.02 | 7.42 | 0.67 |
| | Greenhouse-Geisser | 62.53 | 1.93 | 32.24 | 3.71 | 0.02 | 0.02 | 7.19 | 0.66 |
| | Huynh-Feldt | 62.53 | 2.00 | 31.26 | 3.71 | 0.02 | 0.02 | 7.42 | 0.67 |
| | Lower-bound | 62.53 | 1.00 | 62.53 | 3.71 | 0.05 | 0.02 | 3.71 | 0.48 |
| Task * Study Groups | Sphericity Assumed | 39.96 | 6 | 6.66 | 0.79 | 0.57 | 0.01 | 4.74 | 0.31 |

Table 4.15 (Continued)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Greenhouse-Geisser | 39.961 | 5.81 | 6.86 | 0.79 | 0.57 | 0.01 | 4.59 | 0.30 |
| | Huynh-Feldt | 39.96 | 6.00 | 6.66 | 0.79 | 0.57 | 0.01 | 4.74 | 0.31 |
| | Lower-bound | 39.96 | 3.00 | 13.32 | 0.79 | 0.50 | 0.01 | 2.37 | 0.21 |
| Error(stage) | Sphericity Assumed | 2089.96 | 248 | 8.42 | | | | | |
| | Greenhouse-Geisser | 2089.96 | 240.45 | 8.69 | | | | | |
| | Huynh-Feldt | 2089.96 | 248.00 | 8.42 | | | | | |
| | Lower-bound | 2089.96 | 124.00 | 16.85 | | | | | |

In our study, we looked at how negative feelings changed over different stages (Table 4.9). The analysis of within-subjects effects revealed a significant main effect of stage on negative feelings, with an F-value ($F(2, 24) = 3.71$) and a significance level ($p = 0.02$). This suggests that the stage of the activity has a statistically significant impact on the negative feelings experienced by participants.

Table 4.16 Descriptive Statistics for Negative Feelings about the tool

| **Descriptive Statistics** | | | | |
|---|---|---|---|---|
| Study Groups | | Mean | Std. Deviation | N |
| NEGATIVE_FEELING_ABOUT_TOOL in Task 1 – Knowing and understanding | My memory | 13.51 | 3.74 | 31 |
| | E-textbook | 13.14 | 4.31 | 27 |
| | Google | 11.81 | 4.57 | 32 |
| | ChatGPT | 12.13 | 5.31 | 38 |
| | Total | 12.60 | 4.57 | 128 |
| NEGATIVE_FEELING_ABOUT_TOOL in Task 2 – Applying | My memory | 13.45 | 4.04 | 31 |
| | E-textbook | 13.44 | 3.45 | 27 |
| | Google | 12.81 | 4.13 | 32 |
| | ChatGPT | 13.39 | 5.01 | 38 |
| | Total | 13.2734 | 4.22 | 128 |
| NEGATIVE_FEELING_ABOUT_TOOL in Task 3 - Synthesizing, Evaluating, and Creating | My memory | 13.32 | 5.13 | 31 |
| | E-textbook | 14.22 | 3.88 | 27 |
| | Google | 13.50 | 4.45 | 32 |
| | ChatGPT | 13.50 | 5.06 | 38 |

Table 4.17 (Continued)

| | | | | |
|---|---|---|---|---|
| | Total | 13.60 | 4.66 | 128 |

Our research revealed that students' judgments of unpleasant feelings connected with various information and technology tools differed throughout stages. As seen in the (Table 4.10) In Task 1, students reported the most unpleasant sensations when relying on their memories (mean = 13.51) and the least when using Google (mean = 11.81). ChatGPT was regarded more positively than control group, but it produced more negative emotions than Google (mean = 12.13).

As the tasks progressed to Task 2, the trend in negative feelings remained relatively consistent for control group (Mean = 13.45) and Google (Mean = 12.81), while e-textbooks (Mean = 13.44) and ChatGPT (Mean = 13.39) closely aligned and indicated a slightly increased perception of negative feelings.

However, in Task 3, the impression of unfavorable feelings toward e-textbooks increased dramatically, with the highest mean of 14.22, indicating that e-textbooks were seen as less desirable at this advanced stage. Both control group and Google increased to mean values of 13.32 and 13.50, respectively, whereas ChatGPT's negative emotion mean remained steady from Task 2 to Task 3.

From the most to the least unfavorable feelings throughout all stages, control group consistently provoked the strongest negative sensations, followed by E-textbooks, which peaked in Task 3. Google and ChatGPT were seen to elicit fewer negative emotions, with ChatGPT being the most constant across stages, indicating its potential as a supporting tool in education. Surprisingly, the total mean of negative sensations grew gradually from Task 1 to Task 3, indicating a general tendency of growing negative feelings as cognitive demands of activities increased. However, the proportionate rise for ChatGPT users from Task 2 to Task 3 was the smallest, highlighting its potential usefulness in assisting students as they tackle more complicated tasks.

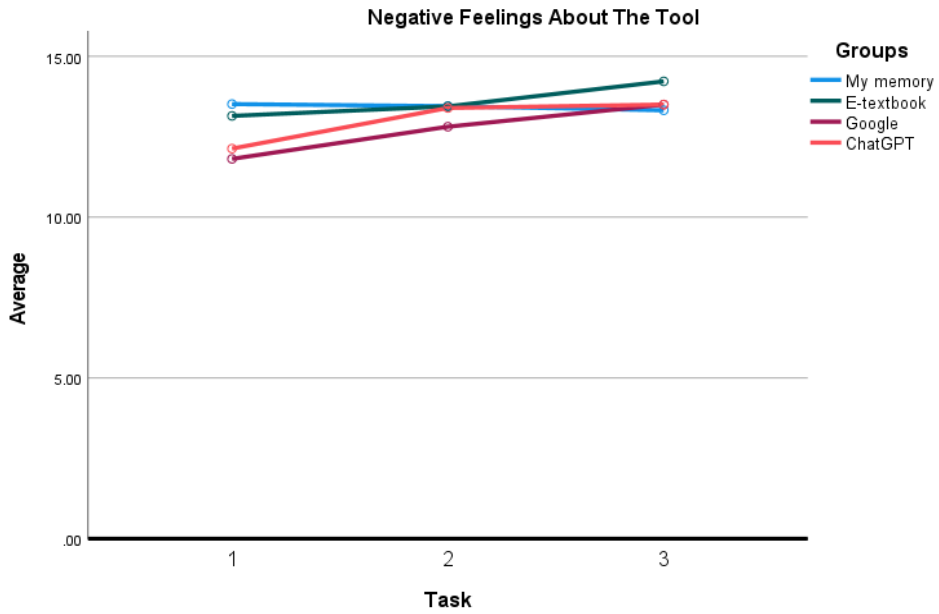


Figure 4.5 Negative Feelings about the tool Trends by Group and Task

Three stages revealed differences in unfavorable perceptions among the groups. Initially, students who relied on 'Google' and 'ChatGPT' reported lower levels of negative feelings as shown in (Figure 4.5), indicating a more positive opinion of these tools, whereas control group and those who relied on 'E-textbooks' reported more negative feelings. From Stage 1 to Task 2, the 'Google' group had a modest increase in unfavorable feelings while maintaining the lowest negative impressions. In contrast, unfavorable attitudes toward the 'E-textbook' group increased significantly, notably in Task 2, showing a rising dissatisfaction with the difficulty of assignments. The 'ChatGPT' group's negative feelings climbed slightly, reflecting the trend shown in the control group, before converging at the conclusion. These trends indicate that students' feelings to task difficulty differ depending on the tool employed, with digital aids like 'Google' and 'ChatGPT' being connected with less negative sentiment than traditional approaches such as memory and e-textbooks. However, the increase in negative feelings across all tools by Task 2 demonstrates that more complex activities may influence students' feelings to the tools, emphasizing the significance of assessing how these tools serve students as tasks grow more difficult.

## 4.6 Retention quiz scores

### 4.6.1 Retention Task 1– Knowing and understanding

We run the analysis for Sphericity Assumed and Wilks' Lambda for Task 1 – Knowing and understanding to ensure that our results are statistically significant Sphericity Assumed gave (F-value = 5.68, p-value = 0.001) and Wilks' Lambda gave (F-value = 5.68, p-value = 0.001) these values indicate that the results are statistically significant as the p-value for Sphericity Assumed and Wilks' Lambda is less than the commonly used threshold of 0.05.

Table 4.18 Descriptive Statistics for Retention Task 1– Knowing and understanding

| **Descriptive Statistics** | | | | |
|---|---|---|---|---|
| Groups | | Mean | Std. Deviation | N |
| Quiz Scores in Task 1 – Knowing and understanding | My memory | 60.2145 | 17.63211 | 31 |
| | E-textbook | 66.4615 | 20.69917 | 27 |
| | Google | 78.9928 | 17.66518 | 32 |
| | ChatGPT | 82.6024 | 17.76440 | 38 |
| | Total | 72.8732 | 20.36855 | 128 |
| Retention Scores in Task 1 – Knowing and understanding | My memory | 59.3162 | 22.05745 | 31 |
| | E-textbook | 68.3062 | 24.88845 | 27 |
| | Google | 66.4860 | 24.62721 | 32 |
| | ChatGPT | 65.5281 | 22.04726 | 38 |
| | Total | 64.8491 | 23.28575 | 128 |

Now according to the table we can observe that The E-textbook group favored with a mean score of 68.31. This is followed by the Google group, which obtained a mean retention score of 66.49. The ChatGPT group had a somewhat lower mean retention score of 65.53. The My memory group again performed poorly, with a mean retention score of 59.32. The average retention score across all groups is 64.85.

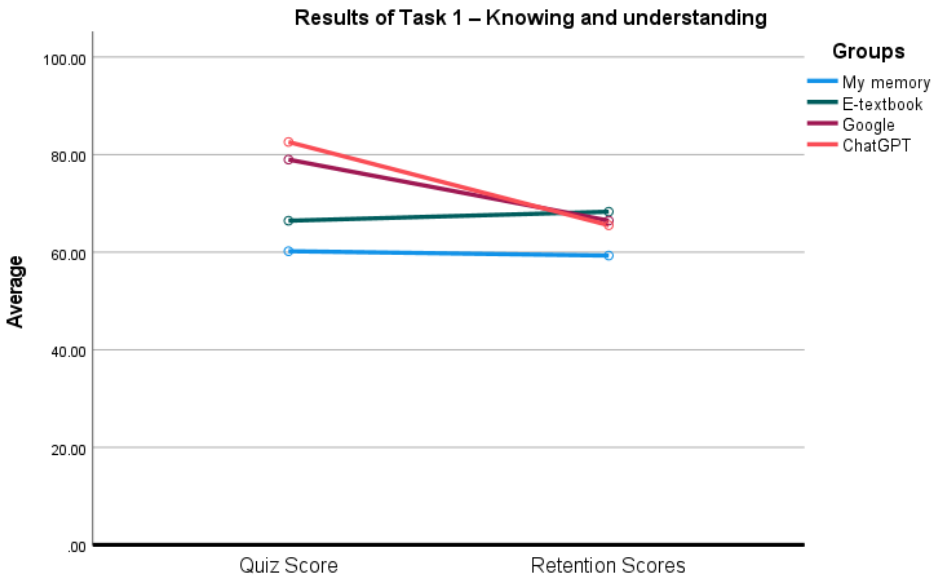


Figure 4.6 Quiz Score and Retention Score Trends among the groups in Task 1

The graph shows students' scores in Task 1. We can see that ChatGPT and Google had the highest Quiz Score values, but it decreased significantly in retention scores. On the other hand, we can see that E-textbook is more beneficial; even though they did not have the highest quiz scores, they had nearly the same score in the retention exam, indicating that E-textbook is slightly more beneficial for retention. My memory group consistently underperformed on both quizzes and retention tests.

### 4.6.2 Retention Task 2 – Applying

We ran the analysis for Sphericity Assumed and Wilks' Lambda for Task 2 * to ensure that our findings were statistically significant. Sphericity Assumed resulted in (F-value = 0.47, p-value = 0.702 ) , And Wilks' Lambda gave (F-value = 0.47, p-value = 0.702). These statistics suggest that the results are not statistically significant, as the p-value for both Sphericity Assumed and Wilks' Lambda exceeds the generally used 0.05 threshold.

Table 4.19 Descriptive Statistics for Retention Task 2 - Applying

| **Descriptive Statistics** | | | | |
|---|---|---|---|---|
| Groups | | Mean | Std. Deviation | N |
| Quiz Scores in Task 2 – Applying | My memory | 47.74 | 19.78 | 31 |
| | E-textbook | 61.85 | 22.36 | 27 |
| | Google | 66.25 | 16.21 | 32 |
| | ChatGPT | 67.89 | 20.02 | 38 |
| | Total | 61.32 | 20.97 | 128 |
| Retention Scores in Task 2 – Applying | My memory | 50.00 | 21.44 | 31 |
| | E-textbook | 63.33 | 22.01 | 27 |
| | Google | 63.75 | 23.10 | 32 |
| | ChatGPT | 64.68 | 22.85 | 38 |
| | Total | 60.61 | 22.95 | 128 |

In the table () we observe that The ChatGPT group had a mean retention score of 64.69. The Google group comes next, with a mean retention score of 63.75. The E-textbook group got a somewhat lower mean retention score (63.33). The My memory group did poorly again, with a mean retention score of 50.00. The average retention score for all groups is 60.61.

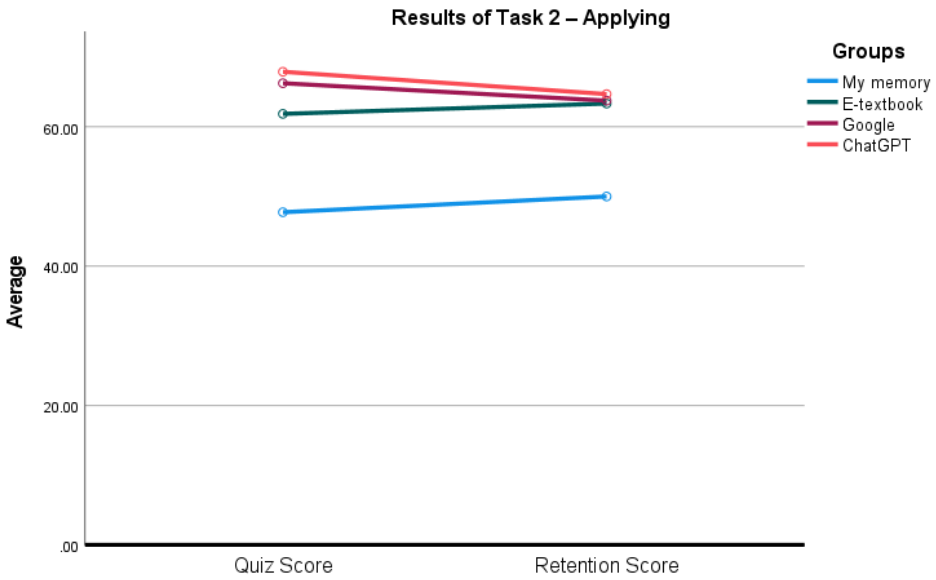


Figure 4.7 Quiz Score and Retention Score Trends among the groups in Task 2

The graph illustrates that, despite the ChatGPT group had the highest quiz scores, their retention scores are not significantly greater than their quiz scores. The E-textbook group performs well overall, with reasonably high retention ratings that outperform other groups. The Google group had a slight drop in retention relative to quiz scores, while the My memory group continuously has the lowest quiz and retention scores. In the end we can see that E-textbooks demonstrate the best overall retention.

### 4.6.3 Retention Task 3 - Synthesizing, Evaluating, and Creating

We ran the analysis for Sphericity Assumed and Wilks' Lambda for Task 3 that our findings were statistically significant. Sphericity Assumed resulted in (F-value = 0.49, p-value = 0.689) , And  And Wilks' Lambda gave (F-value = 0.49, p-value = 0.689). These statistics suggest that the results are not statistically significant, as the p-value for both Sphericity Assumed and Wilks' Lambda exceeds the generally used 0.05 threshold.

Table 4.20 Descriptive Statistics for Retention Task 3 Synthesizing, Evaluating, and Creating

| **Descriptive Statistics** | | | | |
|---|---|---|---|---|
| Groups | | Mean | Std. Deviation | N |
| QUIZ in Task 3 - Synthesizing, Evaluating, and Creating | My memory | 46.0071 | 14.70787 | 31 |
| | E-textbook | 41.7011 | 14.64618 | 27 |
| | Google | 41.2228 | 15.04379 | 32 |
| | ChatGPT | 41.1579 | 14.77343 | 38 |
| | Total | 42.4631 | 14.76245 | 128 |
| RETENTION in Task 3 - Synthesizing, Evaluating, and Creating | My memory | 40.2207 | 15.18951 | 31 |
| | E-textbook | 40.0511 | 20.23082 | 27 |
| | Google | 36.5590 | 14.88974 | 32 |
| | ChatGPT | 34.7113 | 18.07748 | 38 |
| | Total | 37.6339 | 17.12715 | 128 |

In terms of retention scores Task 3 we can see in the table () that the My memory group again had the highest mean score of 40.22, closely followed by the E-textbook group with a mean score of 40.05. The Google group had a mean retention score of 36.56, while the ChatGPT group had the lowest mean retention score of 34.71. The overall mean retention score across all groups was 37.63.

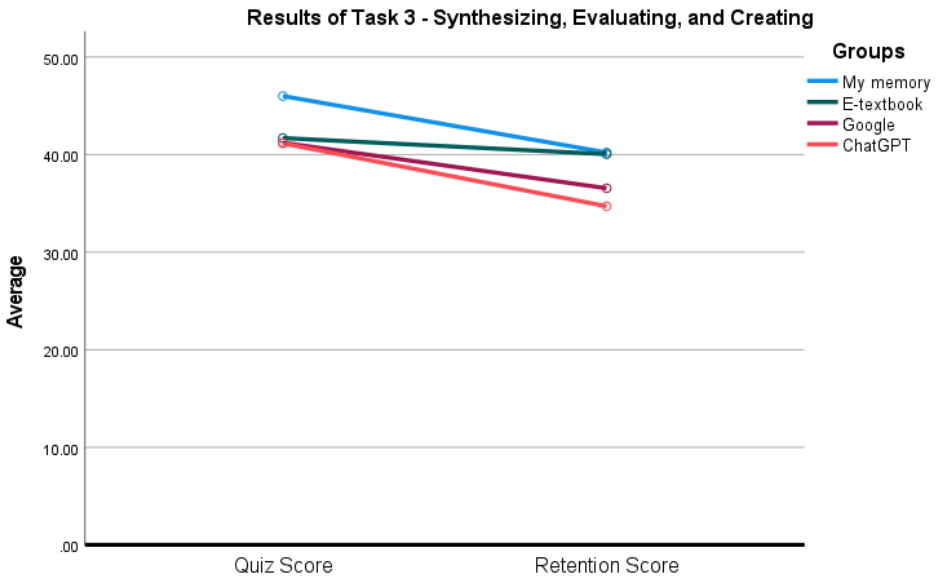


Figure 4.8 Quiz Score and Retention Score Trends among the groups in task 3

The graph shows that the My memory group had highest scores on both the quiz and retention for Task 3. The E-textbook group does well, with retention scores that are comparable to their quiz results. The Google group has similar quiz scores but significantly worse retention scores. Notably, the ChatGPT group performs the worst in terms of retention, earning the lowest score of any group. This implies that for the task involving synthesizing, evaluating, and creating, using ChatGPT resulted in the lowest retention, emphasizing a substantial disadvantage when compared to other informational tools

# CHAPTER 5

# DISCUSSION AND CONCLUSION

## 5.1 INTRODUCTION

The discussion chapter explores the study's findings and contextualizes them within the current research, with a focus on how ChatGPT affects learning. This includes looking at increases in learning efficacy, increased student engagement and motivation, and the emotional and cognitive responses evoked by the employment of AI. Additionally, the chapter addresses practical and ethical problems, emphasizing the importance of these findings for educational practice and policy.

The chapter concludes with a discussion of the study's main findings and their implications for the educational landscape and future research. It offers concrete ideas for educators and educational institutions on how to effectively integrate ChatGVP into teaching approaches. Furthermore, it acknowledges the study's shortcomings and proposes directions for future research. This section emphasizes the overall significance of the research by demonstrating the revolutionary potential of AI technologies such as ChatGPT to improve educational practices and elevate student learning experiences.

## 5.2 DISCUSSION

The major goal of this study was to assess the impact of ChatGPT, a cutting-edge AI conversational agent, on the learning results, engagement, and emotional responses of students in higher education. The study's specific goal was to evaluate the effectiveness of ChatGPT to standard educational tools like e-textbooks and search engines, focusing on how these tools affect students' academic performance, feelings, and retention.

The study's key findings demonstrate that ChatGPT greatly improves learning. Students who used ChatGPT displayed better problem-solving skills and higher levels of interest than those who used traditional methods. Furthermore, ChatGPT users reported favorable emotional responses, such as increased interest and enjoyment in learning activities. However, the study revealed practical and ethical issues, such as potential overreliance on AI and worries about academic integrity.

Discussing these findings in the context of previous research is critical for gaining a thorough grasp of the consequences of using AI tools into education. Previous research has identified comparable benefits and concerns connected with the usage of ChatGPT and other AI technologies in educational contexts. For example, the enhanced learning results and user satisfaction shown in studies [11] and [12] are most likely due to the interactive and adaptive character of AI tools like ChatGPT, which can tailor responses to individual learning styles and immediate informational demands. This individualized interaction improves the learning experience by giving relevant information in an engaging manner, resulting in higher satisfaction and more effective learning.

Moreover, [13] underlines the beneficial effects of AI tools on student motivation. This can be due to the dynamic engagement with AI, which differs from the static character of traditional instructional tools. AI systems frequently provide quick feedback and interact in a conversational manner, making the learning experience more engaging and less intimidating for students. This engaging engagement can boost motivation by making the learning experience more meaningful and pleasurable.

### 5.2.1 Effectiveness of ChatGPT in Enhancing Learning

Our findings show that ChatGPT greatly improves learning effectiveness among higher education students. AI technology offers individualized learning experiences, immediate feedback, and interesting interactions that all contribute to better academic performance [26]. Students who used ChatGPT had better problem-solving skills and retention rates than those who used traditional educational tools like e-textbooks and search engines.

The findings of this study are consistent with the results of [16], which emphasizes ChatGPT's potential to inspire creativity and facilitate the production of unique

solutions. This study shows that ChatGPT not only improves basic learning activities, but it also promotes higher order thinking skills like analysis and synthesis, which are critical for academic achievement.

Furthermore, the comparative study [17] backs up our findings by demonstrating that ChatGPT users spend less time searching for information and rate the quality of information higher. This increased speed in information retrieval leads to more concentrated and effective learning sessions. Users using ChatGPT reported higher levels of enjoyment and perceived utility, which are important elements for enhancing the quality of study sessions.

Additionally, [14] makes an important comparison by investigating how digital tools improve learning. This study discovered that digital textbooks, particularly those with interactive elements, boost student comprehension and performance. ChatGPT, like digital textbooks, offers an interactive and adaptive learning environment. ChatGPT, on the other hand, takes things a step further by engaging in real-time discourse, which allows for dynamic and individualized responses to student inquiries, promoting a deeper comprehension of the subject.

### 5.2.2 Impact on Student Interest

The influence of ChatGPT on student perception related to task interestingness was a major focus of our research. Our data show that the adoption of ChatGPT significantly increased student perception related to task interestingness. Students noted that ChatGPT's interactive and responsive design made learning more enjoyable and less monotonous. The AI tool's fast feedback and individualized guidance increased their interest in finishing assignments.

Our observations are consistent with the findings of [22]. This study found that AI-enhanced learning technologies, such as ChatGPT, improve student perception related to task interestingness by offering personalized help and interesting learning experiences. Both studies found that ChatGPT's individualized feedback and adaptive learning routes increased students' intrinsic interest, resulting in higher levels of interest.

Furthermore, [23] confirms our findings by indicating that students see ChatGPT as a useful and interesting learning partner. The theme analysis revealed that students valued the conversational contact with ChatGPT, which made learning more dynamic and less isolating. This study also discovered that students were more likely to engage in difficult tasks when utilizing ChatGPT, since the AI provided a sense of support and direction throughout the learning process.

While our study and the supporting studies show that ChatGPT has a favorable influence on perception related to task interestingness, there are a few differences to note. For example, our study found that some students initially battled with over-reliance on ChatGPT, expecting the AI to deliver answers rather than guiding them to find solutions on their own. This issue was less evident in [23], which could be attributed to variations in how the AI tool was integrated into the learning environment.

One possible explanation for these discrepancies is the setting in which ChatGPT is employed. In our study, ChatGPT was used in tasks, some of which needed higher-order thinking skills. This expanded application may have presented more substantial issues in balancing reliance on AI with individual problem-solving. In contrast, the research listed may have concentrated on more defined or regulated use cases, resulting in less problems with over-reliance.
To summarize, ChatGPT has a significant positive influence on student involvement and motivation, as demonstrated by our research and supported by current literature. The interactive, supportive, and adaptable character of ChatGPT increases students' intrinsic motivation and engagement. However, the risk of over-reliance emphasizes the importance of well planned integration tactics to ensure that students develop independent learning skills while using AI tools.

### 5.2.3 Emotional and Cognitive Responses to AI Tools

Our study looked into students' emotional and cognitive reactions to ChatGPT and found a variety of good responses. Students reported feeling more interested and less anxious when utilizing ChatGPT instead of traditional learning tools. The AI's capacity to provide fast feedback and individualized help was identified as a major contributor to these favorable emotional responses. Many students reported that the interactive feature of ChatGPT made learning more enjoyable and lowered anxiety levels,

increasing their cognitive engagement.
These results are similar to those from [23]. This study found that the employment of modern technological tools in education can improve students' psychological well-being by lowering anxiety and enhancing perception related to task interestingness. Like our findings, the study discovered that immediate and tailored feedback from AI technologies makes students feel more supported and confident in their learning process, fostering a good learning environment.

Furthermore, [23] adds context by covering both enthusiasm and skepticism about the application of ChatGPT in educational settings. While our study showed primarily favorable emotional responses, this evaluation demonstrates that some students and educators are suspicious about the dependability and consistency of AI-generated responses. Despite these problems, the assessment highlights ChatGPT's potential to improve cognitive perception related to task interestingness through its interactive features and extensive knowledge base.

In our study, using ChatGPT significantly increased students' cognitive engagement. However, there is a concern raised in our research was the risk of over-reliance on AI tools. Some students expressed concern that they would become overly reliant on ChatGPT for solutions, which could impede the development of autonomous problem-solving abilities. This problem was less prevalent in other studies, but it remains an important factor to consider in the efficient integration of AI in education.

### 5.2.4 Comparative Analysis of AI and Traditional Tools

When doing a comparative examination of ChatGPT and traditional educational tools, it is critical to analyze the many strengths and disadvantages of each method. This analysis uses findings from several studies, including [28] and [25], to provide a comprehensive understanding of how AI tools like ChatGPT compare to traditional methods in terms of user satisfaction, perception related to task interestingness, and learning outcomes.

#### 5.2.4.1 Strengths of AI Tools

One of ChatGPT's key assets is its ability to deliver tailored, quick feedback and support. The study [28] finds that ChatGPT users are more satisfied due to the AI's

capacity to offer instant, context-aware responses. This fast feedback loop promotes student perception related to task interestingness by allowing users to obtain responses to their questions in real time, which is sometimes not achievable with older tools.

Furthermore, ChatGPT's interactive aspect promotes a more interesting learning experience. The article [25] concludes that AI chatbots improve student perception related to task interestingness by providing a conversational interface that simulates human interaction. This interaction enables students to dig deeper into topics and promotes active learning, which is a big advantage over static resources such as textbooks.

Another remarkable feature of ChatGPT is its adaptability to different learning styles and needs. [24] emphasizes the ability of AI tools to adjust learning experiences based on the user's choices and performance, providing personalized help that traditional techniques cannot match. This customization is critical for meeting varied learning requirements and ensuring that all students benefit from the instructional content.

#### 5.2.4.2 Weaknesses of AI Tools

Despite their strengths, AI technologies such as ChatGPT have some drawbacks. One major concern is the possibility of misinformation. The paper [23] states that, while ChatGPT can deliver accurate information, it is also prone to producing wrong or misleading responses. This issue needs critical examination and cross-verification by students, which may be disadvantageous when compared to traditional textbooks that have undergone rigorous editorial processes.

Another drawback is the possibility of overreliance on AI. As discussed in [18], there is concern that students will become overly reliant on AI tools for answers, potentially jeopardizing the development of independent critical thinking and problem-solving abilities. Educators must carefully balance the usage of AI tools such that they supplement rather than replace traditional learning techniques.

### 5.2.5 Comparative Insights

When comparing ChatGPT to traditional resources like textbooks and search engines, some noteworthy findings emerge. Traditional resources, which is textbooks, provide well-structured, peer-reviewed content that is credible and easy to reference. However,

they lack the engagement and fast feedback that ChatGPT offers. Search engines such as Google can swiftly retrieve large volumes of information, but they lack the conversational, context-specific support that ChatGPT provides.

In terms of user happiness, the study [28] discovered that people prefer ChatGPT due to its simplicity of use and response quality. Similarly, [25] shows that the engaging nature of AI chatbots results in better levels of student interest and satisfaction when compared to traditional technologies.

### 5.2.6 Integration of AI Tools in Educational Practices

Our findings have important implications for the integration of AI tools such as ChatGPT into educational practices. By effectively adding ChatGPT into their teaching practices, educators may increase student perception related to task interestingness, improve learning outcomes, and build a more dynamic and individualized learning environment.

To effectively integrate ChatGPT, instructors should explore introducing the AI technology into a variety of curriculum areas. For example, ChatGPT can be used to deliver immediate feedback on assignments, assist students with problem-solving exercises, and facilitate discussions on complex topics. [24] discusses how AI technologies can be used to provide personalized learning experiences that are tailored to the specific needs and learning styles of students. This method can considerably improve the entire learning experience by making it more personalized and sensitive to student needs.

Policy recommendations based on our study's findings underline the importance of comprehensive educator training programs to guarantee that they are well-prepared to use AI tools into their teaching practices. Educators must be trained not only in the technical aspects of employing ChatGPT, but also in devising activities and evaluations that fully utilize the AI's capabilities. [16] emphasizes the necessity of professional development for educators in adjusting to new digital tools, which is also applicable to the integration of AI technologies such as ChatGPT.

Furthermore, institutions must establish explicit standards and procedures for the ethical use of AI tools in education. These regulations should address concerns about

data privacy, academic integrity, and the proper use of AI-generated content. [23]contends that existing evaluation methods must be examined and revised to reflect the possibilities and problems posed by AI tools. Taking a more holistic approach to assessment allows instructors to ensure that students are graded fairly and fully.

In addition to these tactics, ongoing evaluation and feedback are critical for successfully integrating AI tools into education. Institutions should routinely analyze the success of ChatGPT and other AI tools in improving learning outcomes and make required changes based on student and educator feedback. [15] emphasizes the necessity of continuous evaluation in understanding the influence of digital tools on learning, which also applies to AI integration.

Finally, cultivating an ethical AI culture is critical for maximizing ChatGPT's benefits while minimizing potential hazards. [21] examines the ethical concerns and potential biases related with AI technologies, highlighting the importance of transparent and responsible AI activities. Educators and institutions must prioritize ethical issues in AI policies and ensure that students understand the ethical consequences of utilizing AI in their learning.

### 5.2.7 Practical and Ethical Challenges

Our study highlighted many practical and ethical obstacles associated with integrating ChatGPT into educational settings. These concerns must be carefully considered to ensure the successful and responsible use of AI tools in education.

One of the most significant practical challenges identified in our research is the potential reliance on AI. Students may become unduly reliant on ChatGPT for answers and assistance, impeding the development of autonomous problem-solving and critical thinking abilities. This issue of dependency was reflected in [18], which discovered that students may rely too much on chatbots, limiting their involvement in active learning. To address this, instructors must create learning activities that encourage students to use AI as a supplement rather than the primary source of information.

Implementation concerns present considerable practical challenges. Integrating ChatGPT into existing educational frameworks involves significant resources, such as technology infrastructure, instructor training, and ongoing student support. [19]

identified similar issues, highlighting the necessity for proper infrastructure and training to realize the benefits of ChatGPT. Institutions must invest in these areas to guarantee that AI tools are seamlessly integrated into their curriculum.

When bringing artificial intelligence into education, ethical considerations are crucial. Academic integrity is a major concern, as ChatGPT may encourage cheating by giving students quick access to answers. [20] outlines solutions for addressing this issue, such as creating examinations to reduce the usage of AI-generated content and deploying plagiarism detection systems. Educators must be cautious when developing evaluations that measure students' comprehension and skills rather than their ability to use AI tools.

Data privacy is another major ethical issue. The usage of ChatGPT requires the gathering and processing of personal data, which raises concerns regarding data security and the ethical use of student information. [21] underlines the significance of data privacy and the need for stringent regulations to safeguard sensitive information. Institutions must develop clear norms and practices to ensure that student data is handled appropriately and transparently.

## 5.3 CONCLUSION

The primary objective of this study was to assess the impact of ChatGPT, a cutting-edge AI conversational agent, on a variety of learning outcomes among higher education students. The study focused on task difficulty, task interest, positive and negative sentiments about the tool, quiz results, and retention rates. The study sought to evaluate the efficacy of ChatGPT to traditional educational resources such as e-textbooks and search engines.

### 5.3.1 Summary of Key Findings

ChatGPT considerably decreased the perceived difficulty of activities across cognitive levels. Students who used ChatGPT felt that tasks were easier than those using Google and PDF documents. ChatGPT's interactive and responsive character significantly increased student engagement and motivation. ChatGPT made learning activities more exciting and enjoyable for students, resulting in increased involvement and excitement. Students who used ChatGPT outperformed their peers on quizzes,

especially in the early phases of learning. Although the performance advantage diminished as tasks got more complicated, ChatGPT still improved initial comprehension and retention of knowledge. ChatGPT resulted in favorable emotional responses from students, who felt less anxious and more supported. However, the study also highlighted significant ethical and practical challenges, such as the dangers of overreliance on AI and worries about academic integrity.

### 5.3.2 Limitations of the Study

While this study provides useful information about the impact of ChatGPT on various educational outcomes, numerous limitations must be addressed. First, the study sample was limited to students from a single university, which may not be indicative of the overall student population. This constraint may limit the findings' generalizability to other educational contexts or demographic groupings. Furthermore, the study was mostly based on self-reported measures of involvement and emotional responses, which are susceptible to biases such as social desirability and recall bias. The experimental design, while rigorous, did not fully replicate real-world learning contexts in which students may use many tools at the same time. Furthermore, the study was rather brief, focused on immediate and short-term effects rather than long-term consequences. These limitations imply that, while the findings are informative, they should be regarded with caution, and that additional study is required to validate and expand on these findings.

### 5.3.3 Future Research Directions

Future research should look into a number of critical areas to expand on the findings of this study. Longitudinal research is required to determine the long-term effects of AI tools such as ChatGPT on learning outcomes, retention, and academic performance. Such research would aid in determining if the early gains noticed are long-term and how they affect overall academic success. Furthermore, future research should incorporate more varied and larger sample sizes to improve the findings' generalizability across other educational contexts and groups. Investigating ChatGPT's influence across disciplines and educational levels, from primary to higher education, could provide a more complete picture of its success. Another major area of research is the combination of AI tools with other educational technology and approaches to

uncover best practices for establishing blended learning environments. Finally, investigating the ethical implications of AI in education, such as data privacy, access equality, and academic integrity, should remain a top goal to ensure that these technologies help all students in a fair and responsible manner.

### 5.3.4 Final Thoughts

The ability of AI tools such as ChatGPT to revolutionize educational methods is enormous. As this study demonstrated, AI can greatly increase student engagement, reduce perceived task complexity, and improve early learning outcomes. However, successfully integrating AI into education necessitates careful consideration of ethical and practical issues. Continued research is required to understand the long-term implications and establish ways that maximize the benefits of AI while limiting any negatives. The deliberate integration of technology in education, supported by strong regulations and inclusive practices, can result in dynamic and individualized learning experiences that prepare students for the challenges of the future. As educators, policymakers, and academics, we must continuously collaborate and innovate to ensure that AI tools are used efficiently and ethically to improve education quality for all students.

## REFERENCES


[1] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education – where are the educators?," *International Journal of Educational Technology in Higher Education*, vol. 16, no. 1, pp. 1-27 Oct. 2019.

[2] I. Červenková, Marek Václavík, and Zuzana Sikorová, "Ways of Using Textbooks and Digital Resources by Teachers and Students in the Lower Secondary Schools," *The New Educational Review*, vol. 71, no. 1, pp. 24–34, Jan. 2023.

[3] Eszter Hargittai, L. Fullerton, E. Menchen-Trevino, and Kristin Yates Thomas, "Trust Online: Young Adults' Evaluation of Web Content," *International journal of communication*, vol. 4, no. 1932–8036, p. 27, Apr. 2010.

[4] A.W. (Tony) Bates, (2022, Augest) ,*Teaching in a Digital Age - Second Edition*, (2nd edition), [online] , Available: https://pressbooks.bccampus.ca/teachinginadigitalagev2/ [March 1, 2024]

[5] J. Paul Gee, "What Video Games Have to Teach Us about Learning and Literacy," *Education + Training*, vol. 46, no. 4, May 2004

[6] R. A. Calvo and S. D'Mello, "Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications," *IEEE Transactions on Affective Computing*, vol. 1, no. 1, pp. 18–37, Jan. 2010.

[7] O. Khassawneh, T. Mohammad, R. Ben-Abdallah, and S. Alabidi, "The Relationship between Emotional Intelligence and Educators' Performance in Higher Education Sector," *Behavioral Sciences*, vol. 12, no. 12, p. 511, Dec. 2022

[8] J. K. M. Ali, M. A. A. Shamsan, T. A. Hezam, and A. A. Q. Mohammed, "Impact of chatgpt on learning motivation," *Journal of English Studies in Arabia Felix*, vol. 2, no. 1, pp. 41–49, Mar. 2023.

[9] K. Seo, J. Tang, I. Roll, S. Fels, and D. Yoon, "The Impact of Artificial Intelligence on Learner–instructor Interaction in Online Learning," *International Journal of Educational Technology in Higher Education*, vol. 18, no. 1, pp. 1–23, Oct. 2021.

[10] B. Tonbuloğlu, "An Evaluation of the Use of Artificial Intelligence Applications in Online Education," *Journal of educational technology and online learning*, vol. 6, no. 4, Sep. 2023.

[11] H. B. Essel, D. Vlachopoulos, A. B. Essuman, and J. O. Amankwa, "ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language models (LLMs)," *Computers and Education: Artificial Intelligence*, vol. 6, no. 100198, p. 100198, Jun. 2024.

[12] M. M. Amin, R. Mao, E. Cambria, and B. W. Schuller, "A Wide Evaluation of ChatGPT on Affective Computing Tasks," *arXiv.org*, Aug. 26, 2023, Available: https://arxiv.org/abs/2308.13911 [March. 1, 2023]

[13] Artem Zadorozhnyy and W. Yee, "ChatGPT and L2 Written Communication: A Game-Changer or Just Another Tool?," *Languages*, vol. 9, no. 1, pp. 5–5, Dec. 2023.

[14] C. K. Lo, "What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature," *Education Sciences*, vol. 13, no. 4, p. 410, Apr. 2023.

[15] D. R. E. Cotton, P. A. Cotton, and J. R. Shipway, "Chatting and cheating: Ensuring academic integrity in the era of ChatGPT," *Innovations in Education and Teaching International*, vol. 61, no. 2, pp. 228–239, Mar. 2023.

[16] A. Haleem, M. Javaid, and R. P. Singh, "An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges," *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, vol. 2, no. 4, p. 100089, Mar. 2022.

[17] M. Mustamin, S. Saleh, Abd. R. Razak, I. Muchtar, and S. Suriyati, "Islamic Educational Values in Local Wisdom Traditional Tradition of Mappogau Sihanua Karampuang Sinjai District | IJORER : International Journal of Recent Educational Research," *journal.ia-education.com*, vol. 4, no. 2, pp 2509-2520, Mar. 2024.

[18] K. Lu, "Can ChatGPT Help College Instructors Generate High-quality Quiz Questions?," *AHEF Open Access*, vol. 70, pp. 311-318 Jan. 2023.

[19] R. Xu, Y. (Katherine) Feng, and H. Chen, "ChatGPT vs. Google: A Comparative Study of Search Performance and User Experience," *arXiv:2307.01135[cs]*, Jul. 2023 Available: https://arxiv.org/abs/2307.01135 [March. 1, 2023]

[20] J. Chen, Z. Zhuo, and J. Lin, "Does ChatGPT Play a Double-Edged Sword Role in the Field of Higher Education? An In-Depth Exploration of the Factors Affecting Student Performance," *Sustainability*, vol. 15, no. 24, pp. 16928–16928, Dec. 2023.

[21] M. Hmoud, H. Swaity, N. Hamad, O. Karram, and W. Daher, "Higher Education Students' Task Motivation in the Generative Artificial Intelligence Context: The Case of ChatGPT," *Information*, vol. 15, no. 1, p. 33, Jan. 2024.

[22] S. Jalil, S. Rafi, T. D. LaToza, K. Moran, and W. Lam, "ChatGPT and Software Testing Education: Promises & Perils," *arXiv:2302.03287 [cs]*, Feb. 2023, Available: https://arxiv.org/abs/2302.03287 [March. 1, 2023]

[23] M. Sullivan, A. Kelly, and P. McLaughlan, "ChatGPT in higher education: Considerations for academic integrity and student learning," *Journal of Applied Learning & Teaching*, vol. 6, no. 1, Mar. 2023.

[24] A. Shoufan, "Exploring Students' Perceptions of ChatGPT: Thematic Analysis and Follow-up Survey," *IEEE Access*, vol. 11, pp. 1–1, Jan. 2023.

[25] C. Leeder and C. Shah, "Practicing Critical Evaluation of Online Sources Improves Student Search Behavior," *The Journal of Academic Librarianship*, vol. 42, no. 4, pp. 459–468, Jul. 2016.

[26] S. Lee, J.-H. Lee, and Y. Jeong, "EXPRESS: The Effects of Digital Textbooks on Students' Academic Performance, Academic Interest, and Learning Skills," *Journal of Marketing Research*, vol. 60, no. 4, pp. 792-811,

Aug. 2023.

[27] C. A. G. da Silva, F. N. Ramos, R. V. de Moraes, and E. L. dos Santos, "ChatGPT: Challenges and Benefits in Software Programming for Higher Education," *Sustainability*, vol. 16, no. 3, p. 1245, Jan. 2024.

[28] T. Karunaratne and A. Adesina, "Is it the new Google: Impact of ChatGPT on Students' Information Search Habits," *European Conference on e-Learning*, vol. 22, no. 1, pp. 147–155, Oct. 2023.

[29] T. Woodland, "ChatGPT for Improving Medical Education: Proceed With Caution," *Mayo Clinic proceedings: Digital Health*, vol. 1, no. 3, pp. 294–295, Sep. 2023.

[30] J. Trixa and K. Kaspar, "Information literacy in the digital age: information sources, evaluation strategies, and perceived teaching competences of pre-service teachers," *Frontiers in psychology*, vol. 15, Mar. 2024.

[31] V. Salido, "Impact of AI-Powered Learning Tools on Student Understanding and Academic Performance," 2023.

[32] G. Sebastian, "Privacy and Data Protection in ChatGPT and Other AI Chatbots: Strategies for Securing User Information," *International Journal of Security and Privacy in Pervasive Computing (IJSPPC)*, vol. 15, no. 1, pp. 1–14, 2023.

[33] A. Afzal, S. Khan, S. Daud, Z. Ahmad, and A. Butt, "Addressing the Digital Divide: Access and Use of Technology in Education," *Spring 2023*, vol. 3, no. 2, pp. 883–895, Jun. 2023.

[34] T. Waltzer, R. L. Cox, and G. D. Heyman, "Testing the Ability of Teachers and Students to Differentiate between Essays Generated by ChatGPT and High School Students," *Human Behavior and Emerging Technologies*, p. e1923981, Jun. 2023.

[35] M. Agbese, R. Mohanani, A. Ali Khan, and P. Abrahamsson, "Ethical Requirements Stack: A framework for implementing ethical requirements of AI in software engineering practices," *Association for Computing MachineryNew YorkNYUnited States*, Jun. 2023.

[36] A. Almusaed, A. Almssad, I. Yitmen, and R. Z. Homod, "Enhancing Student Engagement: Harnessing 'AIED''s Power in Hybrid Education—A Review Analysis," *Education Sciences*, vol. 13, no. 7, p. 632, Jul. 2023.

[37] R. F. Kizilcec, “To Advance AI Use in Education, Focus on Understanding Educators,” *International Journal of Artificial Intelligence in Education*, vol. 34, pp. 12–19, Jun. 2023.

[38] A. Field and G. Hole, “Experimental Designs” in *How to Design and Report Experiments*, 1st ed. London: Sage Publications, 2013, pp. 54-104.

[39] D. R. Krathwohl, “A revision of bloom’s taxonomy: An overview,” *Theory Into Practice*, vol. 41, no. 4, pp. 212–218, Nov. 2002.

[40] A. Field, “GLM 4: Repeated-measures designs” in *Discovering Statistics Using IBM SPSS Statistics*, 5th ed. Los Angeles: Sage Publications, 2018, pp. 840-886.