Enhancing Adaptive Personalized Learning Interfaces with Generative AI for Individuals with ADHD

P.R. Lakshika
L.D.R.N. Gunawardana
M.D.M.N. Perera
2025

Enhancing Adaptive Personalized Learning Interfaces with Generative AI for Individuals with ADHD

P.R. Lakshika
Index No: 20020619

L.D.R.N. Gunawardana
Index No: 20020432

M.D.M.N. Perera
Index No: 20020742

Supervisor: Dr. Chaminda Ranasinghe
Co-supervisor: Dr. Kasun Karunanayaka

May 2025

Submitted in partial fulfillment of the requirements of the B.Sc. (Honours) Bachelor of Science in Information Systems Final Year Project

Declaration

I certify that this dissertation does not incorporate, without acknowledgement, any material previously submitted for a degree or diploma in any university and to the best of my knowledge and belief, it does not contain any material previously published or written by another person or myself except where due reference is made in the text. I also hereby give consent for my dissertation, if accepted, be made available for photocopying and for interlibrary loans, and for the title and abstract to be made available to outside organizations.

Candidate Name: L.D.R.N Gunawardana

.... Date: 30/06/2025

Signature of Candidate

I certify that this dissertation does not incorporate, without acknowledgement, any material previously submitted for a degree or diploma in any university and to the best of my knowledge and belief, it does not contain any material previously published or written by another person or myself except where due reference is made in the text. I also hereby give consent for my dissertation, if accepted, be made available for photocopying and for interlibrary loans, and for the title and abstract to be made available to outside organizations.

Candidate Name: P.R.Lakshika

.... Date: 30/06/2025

Signature of Candidate

I certify that this dissertation does not incorporate, without acknowledgement, any material previously submitted for a degree or diploma in any university and to the best of my knowledge and belief, it does not contain any material previously published or written by another person or myself except where due reference is made in the text. I also hereby give consent for my dissertation, if accepted, be made available for photocopying and for interlibrary loans, and for the title and abstract to be made available to outside organizations.

Candidate Name: M.D.M.N Perera

.... Date: 30/06/2025

Signature of Candidate

This is to certify that this dissertation is based on the work of

Ms. P.R.Lakshika
Ms. L.D.R.N Gunawardana
Ms. M.D.M.N Perera

under my supervision. The thesis has been prepared according to the format stipulated and is of acceptable standard.

Supervisor's Name: Dr. Chaminda Ranasinghe

.... Date: 30/06/2025
Signature of Supervisor

Principle/Co- Supervisor's Name: Dr. Kasun Karunanayaka

.... Date: 30/06/2025
Signature of Supervisor

Abstract

Students with Attention Deficit Hyperactivity Disorder (ADHD) face unique challenges in educational environments struggling with their impulsivity, hyperactivity, and inattention issues. Traditional learning platforms often fail to accommodate the specific requirements of ADHD learners, highlighting the need of developing personalized ADHD friendly educational platforms that enhance the learning experience of ADHD individuals by creating an equitable opportunity for them to excel in academics. To bridge this research gap we conducted a study exploring the potential of Generative AI to revolutionize personalized learning interfaces for students with ADHD, aged 18–30, who have prior experience with online learning platforms.

This research aims to develop an innovative learning platform tailored for individuals with ADHD improving their encouragement and overall educational outcomes. By applying principles of Human-Computer Interaction (HCI), we compare three distinct learning management systems (LMS): a conventional LMS used by the general population, a manually optimized LMS designed specifically for ADHD users, and an AI-generated LMS customized to ADHD needs through enhanced prompting techniques. This study involves 22 ADHD students whose feedback on each LMS informs our analysis, allowing us to assess how AI-driven customization can adapt interfaces to diverse cognitive and learning styles.

The evaluation employs multiple approaches to evaluate the usability and effectiveness of each experiment, including the System Usability Scale (SUS), User Experience Questionnaire (UEQ), task analysis, and testing conducted with UI/UX professionals and HCI practitioners, while also evaluating user learning outcomes and engagement levels. The findings reveal contrasting user preferences, with the AI-generated Learning Management System preferred for usability, while the manual system is favoured for effectiveness. The findings indicate that although AI-generated user interfaces enhances usability, human-designed user interfaces are essential for educational effectiveness. The study promotes hybrid methodologies that combine AI efficiency with educator-led instructional design, providing practical insights for user interface developers and institutions focused on achieving a balanced user experience and learning outcomes.

Acknowledgement

We would like to express our sincere gratitude to everyone whose invaluable contributions have been highly important to the completion of this thesis. The journey would not have been possible without their constant encouragement, support, and guidance.

First and foremost, we would like to express our thanks to Dr.Chaminda Ranasinghe, our supervisor, for his exceptional guidance, insightful feedback and unwavering support during this research. We sincerely appreciate the time and effort he dedicated to assisting us in improving our ideas and approaches, as his expertise and commitment have been pivotal in determining the direction of this study.

Furthermore, we are extremely grateful to Dr.Kasun Karunanayaka, our Co-supervisor, for his constructive feedback and invaluable support, which significantly improved the scope and quality of this study. His expertise in the human-computer interaction and research field offered us significant insights that were extremely important to this project's success.

We sincerely appreciate the guidance and assistance provided by Prof. K.P. Hewagamage, Dr.Dilrukshi Gamage and Mrs. Sanduni Thrimahavithana during this endeavor. Further we are extremely thankful for all the lecturers of UCSC who assisted in enhancing our knowledge for their continuous guidance and support.

Additionally, we would like to acknowledge the researchers and developers whose work in personalized learning interfaces and generative AI provided the foundation for this study. We are extremely grateful for their scholarly contributions, and their pioneering work in the field that has inspired and guided our research. We are extremely thankful for the participants who volunteered their time and shared their experiences, which allowed us to test and improve our learning interfaces.

We want to sincerely thank our family and friends for their unwavering encouragement, patience, and support during this journey. Their faith in us kept us motivated during challenging times, and we will always be appreciative of their unconditional affection and understanding.

Finally, we would like to extend our sincere gratitude to everyone who has supported us in any way for your tremendous assistance.

Table of Contents

List of Figures

List of Tables

List of Acronyms

ADHD Attention Deficit Hyperactivity Disorder

AI Artificial Intelligence

ASRS Adult ADHD Self-Report Scale

ATM Automated Teller Machine

CSS Cascading Style Sheets

DSR Design Science Research

DSRM Design Science Research Methodology

ERC Ethics Review Committee

GenAI Generative AI

HCI Human Computer Interaction

IT Information Technology

LLM Large Language Model

LMS Learning Management System

MHA Mental Health America

SUS System Usability Scale

UCSC University of Colombo School of Computing

UCSCIRB University of Colombo School of Computing Institutional Review Board

UEQ User Experience Questionnaire

UI User Interface

UX User Experience

WCAG Web Content Accessibility Guidelines

WHO World Health Organization

Chapter 1

Introduction

According to Furman [16], Attention Deficit Hyperactivity Disorder (ADHD) is a common neurodevelopmental disorder, which affects children's ability to control their spontaneous reactions, including movement, speech, and attentiveness.ADHD implies a distinct neuro-divergent method of the human brain's function which differs from neurotypical people. It usually appears in childhood and persists into adulthood [51]. ADHD can be identified with the persistent, and pervasive symptoms of inattention, hyperactivity, and impulsivity. Erskine et al. [14] state that Long-term ADHD is associated with a high risk of mental disorders, interpersonal issues, and academic setbacks.

A significant number of children with ADHD tend to have particular learning disabilities, which may lead to difficulties in reading, writing, and mathematics [3]. In that case, the specific learning disability and ADHD need to be addressed independently of one another. According to the growing prevalence of ADHD diagnoses and the unique challenges encountered by people with ADHD, there is an urgent need to improve specific educational tools that enable personalized learning for them.Different studies have been conducted to improve personalized learning for neurodivergent students, there is a lack of studies in the area of enhancing personalized learning interfaces for ADHD individuals. Although Gkora [18] has been conducted similar studies to support primary education of ADHD students, there is still a significant lack of research in the area of addressing the specific learning requirements of adult ADHD individuals.

With the rapid enhancements of Information Technology (IT) and Artificial Intelligence (AI), especially in the field of generative AI, there are optimal approaches for achieving such specific learning needs. It has been already proved by the studies that refinement of AI tools like ChatGPT can be used for ADHD therapy in a more effective manner [8]. Through the study, Ivica Pesovski [23] which explores a tool for generating personalized learning materials using Large Language Models (LLMs) and OpenAI's API, it implies that generative AI can be used in personalized education to create tailored educational content for students. Although most of the studies have been focused on customization of the learning content, personalization of the learning interfaces area still remains as unexplored. Even though many commercial applications and open-source frameworks have been designed for user interface generation using AI, there is a lack of specific research related to user interface generation with AI models for enhancing personalized education for individuals with needs.

Our research focuses on utilizing generative AI to design ADHD-friendly user interfaces tailored to the needs of ADHD students. Through this approach, we aim to understand the specific requirements and preferences of ADHD learners and create User Interface (UI) designs that enhance their engagement, focus, and learning effectiveness. This research aims to bridge the gap between conventional educational platforms and the specialized needs of ADHD students by developing, UIs based on user feedback from ADHD and non-ADHD individuals, Human Computer Interaction (HCI) heuristics, and Generative AI. Our goal is to provide a learning environment that is both accessible and engaging, ultimately supporting better academic outcomes and learning experiences for ADHD students.

1.1 Problem Statement

Students with Attention Deficit Hyperactivity Disorder (ADHD) encounter unique challenges in educational settings, including difficulties with impulsivity, hyperactivity, and inattention. Even though there are many conventional online learning platforms available, the majority are unable to meet the unique behavioral and cognitive requirements of ADHD students. A clear research gap can be identified which is still unexplored that is addressing the special learning abilities of ADHD individuals through exploring personalized learning platforms for them. Although some research such as Gkora [18] has been conducted for supporting primary education of ADHD students, still there is a significant lack of research in the area of addressing the specific learning requirements of adult ADHD individuals.

With technological advancements, although most of the current research emphasizes gamification, AR, VR, and sensory tools, they do not directly address ADHD learners' needs within personalized learning environments. Generative AI shows immense potential, yet its application for creating ADHD-friendly Learning Management System (LMS) interfaces remains largely unexplored. When applying Generative AI in personalized learning, most of the studies like Ivica Pesovski [23] which explores a tool for generating personalized learning materials using LLMs, focus on customization of learning content but not on generating personalized learning interfaces. These research gaps emphasize the significant need for optimized learning interfaces suitable for adult ADHD students to excel in their academics.

1.2 Research Questions

The purpose of this research project is to undertake a comprehensive exploration into the inquiries posed by the following research questions:

1.2.1 Primary Research Question

- **How can AI-driven customization improve the effectiveness and usability of LMS for individuals with ADHD compared to conventional and manually optimized LMS platforms?**
  The primary focus of our research is developing three distinct prototypes tailored for ADHD students, such as a conventional LMS used by the general population, a manually optimized LMS designed specifically for ADHD users, and an AI-generated LMS customized to ADHD needs through enhanced prompting techniques and conducting a comparative analysis to explore how AI-driven customization improves the effectiveness and usability of LMS comparing to the other two experiments of conventional and manually optimized LMS. The intention is to explore the best learning experience for the ADHD individuals to enhance their educational outcomes, utilizing the potential of generative AI.

1.2.2 Secondary Research Questions

- **How can user data be effectively used to evaluate the usability of conventional LMS to identify limitations and areas for improvement?**
  The user feedback and data would be helpful to identify the challenges they are facing in the current conventional learning environments, and it would be important to make improvements, tailoring them to facilitate unique user requirements. Through this study we intend to identify the limitations of existing conventional LMS and improve them based on gathering user data.

- **What improvements are necessary in personalized learning systems to support ADHD students?**
  With this we intend to study the existing personalized learning platforms and their limitations to address the unique learning requirements of ADHD students. Through this, we expect to study how these existing learning platforms could improve in terms of HCI standards, UI/UX best practices, and ADHD-friendly design guidelines and how these improvements affect students with ADHD.

- **How can prompting strategies in generative AI be optimized to create user interfaces that enhance engagement and usability for ADHD learners?**
  We will study how prompting techniques in generative AI can be utilized to generate personalized learning interfaces for ADHD individuals by testing different generative AI models with different prompts, following advanced prompting techniques to derive an optimal learning experience for ADHD learners.

- **How can the impact of GenAI-driven personalized learning interfaces on academic performance and usability for individuals with ADHD be assessed?**
  We will evaluate the usability and effectiveness of each prototype, and through them the impact of Generative AI enhanced personalized learning interfaces for the academic performance of ADHD individuals can be evaluated by gathering data from user feedback and conducting qualitative and quantitative analysis. Through that we intend to explore how generative AI can be utilized to enhance the overall educational outcomes of ADHD students.

1.3 Research Motivation

ADHD is a substantial challenge for affected individuals, especially in educational environments where sustained focus and engagement are crucial. ADHD significantly impacts individuals' educational experiences due to their impulsivity, hyperactivity, and inattention issues [3]. Despite the growing prevalence of ADHD diagnoses and the critical need for non-pharmacological educational support, the current conventional educational platforms have become unable to address the difficulties faced by ADHD learners such as managing distractions, maintaining focus, and fostering consistent engagement. Although AI has been used in learning content customization, its potential in personalized UI generation is still unexplored, especially for neurodivergent populations like ADHD.

Our study aims to fill these voids by utilizing the capabilities of generative AI, creating an interactive and personalized learning space that helps students with ADHD improve their learning outcomes and enhance their engagement. By leveraging Generative AI, we intend to reshape the educational experience for ADHD students by designing user interfaces that are friendly to their needs and accommodate various cognitive and behavioral challenges. Additionally, we will investigate the best learning solutions for ADHD students by comparing AI-enhanced user interfaces with those that are manually created and traditional designs. Through insights gained from literature reviews and collecting feedback from students with ADHD, we strive to offer an inclusive, efficient, and empowering educational experience tailored to their unique requirements. Our objective is to establish a supportive and engaging learning atmosphere that provides ADHD students with equal chances to excel academically.

1.4 Goals and Objectives

1.4.1 Goals

The goal of the research is to enhance the learning experience of individuals with ADHD by creating an equitable opportunity for them to excel in academics. Conventional educational environments often present unique challenges for ADHD students, but by exploring personalized learning platforms with ADHD-friendly user interfaces specifically designed to address their unique needs, we can empower them to succeed academically. Through that the ADHD individuals would get the opportunity to have enhanced learning experiences with structured, distraction-free, engaging, and manageable learning environments improving their focus and engagement levels which help to leverage their overall academic performance and reduced dropout rates.

1.4.2 Objectives

- Identifying the specific learning needs and challenges faced by students with ADHD in learning environments
- To develop a prototype of a generative AI-enhanced personalized learning interfaces
- Study the prototype's usability and effectiveness with ADHD individuals through user testing, and feedback gathering and analyse them using qualitative and quantitative methods

1.5 Research Approach

The research followed a deductive approach, and was conducted in 6 main phases, aligned with the Design Science Research Methodology (DSRM) [41]. This study started with a theoretical understanding that ADHD students encounter unique challenges in their conventional learning environments. Based on this we constructed the hypothesis that AI-driven personalized learning interfaces will improve learning outcomes and engagement for ADHD learners. To evaluate this hypothesis, the research involved a comparative analysis of AI-driven interfaces with conventional and manually optimized interfaces, evaluating their usability and effectiveness through systematic evaluations. To evaluate the user interfaces, mixed-methods were used including both qualitative and quantitative data analysis. Insights gained from the evaluation supported the verification of research findings and demonstrating the effectiveness of the artifact in addressing the identified problem.

1.6 Limitations and Scope

The research will be conducted through comparing three distinct LMS that are developed by applying principles of HCI and ADHD design guidelines: a conventional LMS used by the general population, a manually optimized LMS designed specifically for ADHD users, and an AI-generated LMS customized to ADHD needs through enhanced prompting techniques.

1.6.1 In Scope

- Exploring design guidelines and UX best practices in developing learning platforms.
- Exploring and clarifying design guidelines and UX features for ADHD users.
- Developing a clickable mock-up as of conventional LMS used by the general population as the "Base LMS" according to HCI heuristics and UX best practices.
- Investigating the accuracy and features of generative AI models in UI generation and selecting the best-fitting models.
- Training AI models with prompt engineering techniques to generate personalized learning interfaces according to the specific requirements of students with ADHD.
- Developing the clickable mock-ups from the generated wireframes/interfaces as the "AI-Generated LMS" customized to ADHD needs through prompting techniques.
- Developing a clickable mock-up with manually personalized learning interfaces specifically designed for ADHD individuals as the "manually optimized LMS"
- Selecting only the low-positive, low-negative, high-negative range ADHD individuals according to Adult ADHD Self-Report Scale (ASRS v1.1)[39], to participate and evaluate the user interfaces.
- Testing with users and evaluating performance
- The evaluation was conducted with a sample of 22 ADHD students within the age range 18-30.
- Focus on using one module/subject as the learning content

1.6.2 Out of Scope

- Deploying the prototype.
- The study won't explore the platform's long-term maintenance, support, or further enhancements based on gathered feedback from users.
- This research will only focus on developing user interface prototypes, no hardware or physical devices will be developed.

- The research will not involve ADHD diagnosis, medication management, therapy sessions, or other clinical techniques, or the creation or validation of diagnostic instruments or methodologies.

1.6.3 Limitations

- The evaluation was conducted with a limited sample of 22 ADHD students within age range 18-30.
- The study won't explore the platform's effect on long-term knowledge retention
- Individuals/Students who use the system need prior knowledge of using an e-learning platform.
- Only one subject module is implemented with the initial level of the prototype development.

1.7 Contribution

This research study makes significant contributions to the field of personalized learning and HCI, particularly in addressing the educational needs of individuals with ADHD as below:

1. **Exploration of Generative AI-enhanced ADHD-Friendly personalized learning interfaces**

   This study explores the usability and efficiency of the novel idea of creating personalized adaptive learning interfaces that are tailored to the unique needs of students with ADHD. The study shows how advanced prompting techniques and user-centered design principles can be combined to create interfaces that improve focus, engagement, and learning outcomes for students with ADHD by utilizing Generative AI, through developing an AI-generated LMS prototype as one of our experiments. This exploration of the AI-generated interface creation with prompting techniques would be helpful as a foundation for future research and development in AI-driven educational tools for neurodivergent individuals.

2. **Comparative Analysis of Conventional, Manually Optimized, and AI-Generated Learning Interfaces for ADHD Students**

   The study provides a comprehensive comparative analysis of three distinct LMSs:

   - A conventional LMS designed for the general population
   - A manually optimized LMS tailored for ADHD users
   - An AI-generated LMS customized to ADHD needs through enhanced prompting techniques

   This study identifies the benefits and drawbacks of these experiments evaluating the usability and effectiveness of each experiment, employing different qualitative and quantitative methods including System Usability Scale (SUS), User Experience Questionnaire (UEQ), Task Analysis and evaluating the learning outcomes and engagement levels of the users. With the use of generative AI, HCI heuristics, and user feedback from ADHD individuals, this research attempts to bridge the gap between traditional educational platforms and the unique needs of ADHD students. Through this study, we aim to find the best educational environment for ADHD students, supporting them to enhance their academic outcomes and learning experiences.

3. **Identification of Key Design Principles for ADHD-Friendly Learning Interfaces**

   Through the identification and validation of key design principles for ADHD-friendly interfaces, this study assists in expanding knowledge on creating inclusive educational technologies. We have identified and evaluated ADHD-friendly design principles such as minimizing distractions through clean and intuitive layouts, incorporating adaptive features to support focus and task completion, and ensuring flexibility to accommodate individual learning preferences, which assist in enhancing the focus and engagement levels of ADHD students. We have identified these design principles through comprehensive literature review, user feedback, expert evaluations, and iterative testing. They would be helpful and offer practical advice for educators, designers, and developers working on inclusive learning platforms in their future studies and developments.

Chapter 2

Background

2.1 ADHD

2.1.1 What is ADHD?

According to Diagnostic and Statistical Manual of Mental Disorders V (DSM-V), mental disorders that manifest during childhood or adolescence can be divided into three main categories which consist of Neurodevelopmental disorders (NDDs), Autism spectrum disorders (ASDs) and Mental health disorders (MHDs) [5]. Attention Deficit/Hyperactivity Disorder (ADHD) is a common Neurodevelopmental disorder that significantly affects people's ability to perform their daily tasks. NDDs have been noted as a major reason for childhood disabilities, impacting not only the affected child but also their loved ones and society as a whole. As stated in Barua et al. [5], it is estimated that between 30% and 50% of kids with ADHD also have co-occurring mental health disorders, like depression. It indicates that NDDs and MHDs often overlap with one another. Such complex combinations of comorbidities between MHDs and NDDs may lead to learning disabilities and negatively impact a person's quality of life. Because of that when an individual is suffering from ADHD along with MHDs it affects them worstly. Therefore an advanced approach is required to enhance the personalized learning for ADHD individuals, considering their cognitive and attentional differences.

Figure 2.1 from Barua et al. [5], indicates an ADHD affected brain compared to a normal brain. An ADHD affected brain often has a slightly smaller overall brain volume compared to a normal brain.

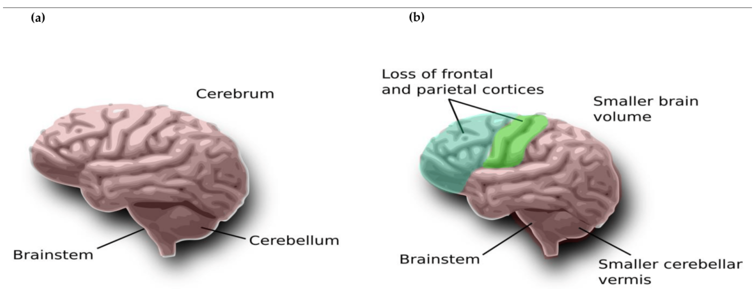


Figure 2.1: (a) Normal brain and (b) ADHD brain with smaller volume [5]

ADHD is a complex disorder influenced by both genetics and environmental factors. The disorder's exact cause is unknown, but studies suggest it is mainly influenced by genetics and the physical environment. As well as non-genetic factors like brain injuries, early birth, maternal smoking, and exposure to environmental toxins like lead are also linked to ADHD. According to the study, Ibrahim et al. [22] which focuses on the ADHD disorders of students between 6-18 years old, there are three forms of ADHD that are prevalent in kids and students.

- Inattentive type : Student is unable to focus or maintain concentration on an assignment or activity. It appears more common in girls to have this disorder.
- Hyperactive-impulsive type : This type of student exhibits excessive behavior and occasionally acts without thinking, but they do not display severe issues with attention. It appears more prevalent among boys to have this disorder.
- Combined type : The learner exhibits impulsivity, hyperactivity, and inattention.

ADHD gender differences are evident in various aspects, including impulsivity, academic achievement, social functioning, parental education, and parental depression. Different research and meta-analysis have been conducted in this area to explore how ADHD impacts gender differences. According to Gaub and Carlson [17] , girls with ADHD are more likely to exhibit intellectual difficulties, less hyperactivity, and fewer externalizing symptoms like aggression or disruptive behaviours compared to boys. In order to determine the prevalence of ADHD, Willcutt established a detailed meta-analysis based on the Diagnostic and Statistical Manual of Mental Disorders IV (DSMIV) [22]. He discovered that there are distinct age and gender disparities in the prevalence of ADHD as displayed below 2.1:

ADHD/subtype	**Age**	**Prevalence**	**Male: female**
Inattentive	6-12	5.1	2.2 : 1
	13-18	5.7	2.0 : 1
Hyperactive - impulsive	6-12	2.9	2.3 : 1
	13-18	1.1	5.5 : 1
Combined type	6-12	3.3	3.6 : 1
	13-18	1.1	5.6 : 1
Total ADHD	6-12	11.4	2.3 : 1
	13-18	8.0	2.4 : 1

Table 2.1: Prevalence of ADHD in children and adolescents

2.1.2 ADHD Severity Scales and Avoiding Associated Risks

When studying about developing a personalized learning approach for the ADHD individuals it is important to consider their severity levels. ADHD manifests differently across each individual, depending on the severity of symptoms and the existence of comorbidities, which may vary from mild to severe levels. Assessing severity is crucial for both reducing possible risks related to high-risk individuals as well as for customizing educational interventions for ADHD students. When conducting research based on the feedback and user evaluations of ADHD students to enhance their learning, the high risk individuals who may suffer with complex comorbidities may not be able to take part in the user evaluations due to associated risks.

Therefore we can use the Adult ADHD Self-Report Scale (ASRS v1.1) [39], which is a commonly used tool to assess ADHD severity of adults, developed in collaboration with the World Health Organization (WHO). It is an 18-item self-report questionnaire designed to assess Attention Deficit Hyperactivity Disorder (ADHD) symptoms in adults (18+). Based on the World Health Organization Composite International Diagnostic Interview (2001), this scale's questions are tailored to address adult symptom manifestation and are in line with both the DSM-IV and DSM-5-TR criteria [25]. Using the symptom checklist of ASRS v1.1 we can evaluate the level of impairment of the ADHD individuals associated with the symptoms and consider the symptom severity according to symptom frequency. By using 0-24 scale scoring method participants can be selected as below.

ADHD Score	**ADHD Range**
0-9	Low negative
10-13	High negative
14-17	Low positive range
18-24	High positive range

Table 2.2: ASRS v1.1 scale scoring

Based on the studies like Kessler et al. [25], this Low Negative (0-9) range ADHD individuals exhibit minimal symptoms with negligible impact on daily functioning. High Negative (10-13) range persons may suffer with noticeable symptoms, but they are still manageable without severe interference in daily activities. In the low positive (14-17) range, symptoms are more prominent, potentially affecting focus, impulsivity, and hyperactivity, requiring targeted support. Individuals belonging to the high positive (18-24) range, are more likely to suffer from severe symptoms often associated with co-occurring other mental health conditions like depression, anxiety, insomnia, panic disorder and facing severe difficulties with managing even daily activities.

Because of this, high positive range individuals above 18-24 are not encouraged to participate in the evaluations of our study. Therefore, we will be selecting the normal ADHD individuals (0-17) here to participate and evaluate the user interfaces to determine what features will impact the most for their educational activities. According to this the normal ADHD individuals can be selected, avoiding the associated risks.

2.1.3 Educational Difficulties Caused by ADHD

Students that suffer from this type of disorder typically have limited attention spans which means the duration of time they can focus on a task and are prone to distractions [22]. Due to ADHD symptoms like inattention, anxiousness, and impulse control negatively affect academic performance and the school environment, children with ADHD frequently struggle in learning environments. As well as they face difficulties in organizing tasks and managing time effectively.

According to Ferguson et al. [15], many university students suffering from ADHD are unaware of their condition, often ashamed and reluctant to seek help, leading to average GPAs significantly lower than those without ADHD. Therefore, undergraduates with ADHD face dual challenges in university life, dealing with disorder symptoms and a new social environment with less parental supervision than previous educational stages. Reluctant to seek help, leading to separateness and stress. They face difficulties in focus and task completion according to impulse control issues. Likewise students with ADHD face more challenges such as new separateness, stress, complex study schedules, and fewer supports, leading to low self-esteem due to high symptoms and low executive functioning. Overall these educational challenges for ADHD students, lead to their lower academic performance and higher dropout rates.

Chapter 3

Literature Review

The below figure 3.1, illustrates the structure of the literature review, organized into the main sections we studied including personalized learning for neurotypical students and ADHD students, existing learning applications for ADHD students, use of GenAI in personalized learning and interface generation.

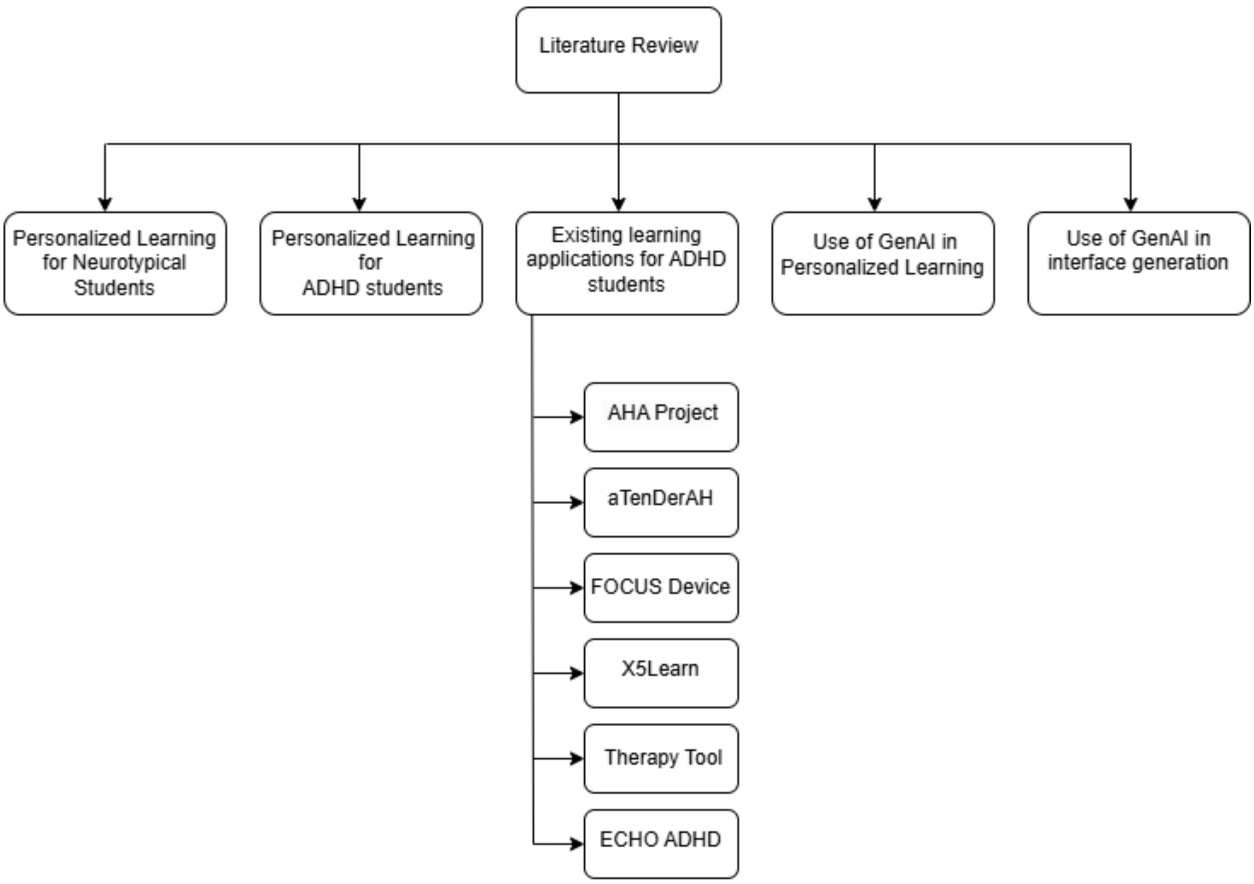


Figure 3.1: Taxonomy Diagram of Literature Review

3.1 Personalized Learning for Neurotypical Students

Personalized learning has become an important educational strategy which enables learners to reach their specific educational needs and abilities by providing adaptive, personalized content based on their potential and skills. In an e-learning platform it is crucial to have personalized interfaces for users ensuring individualization and differentiation because the student group we are focusing on is heterogeneous[50]. When applying personalization it is important to identify the learner first based on learner's intelligence according to Gardner's "Multiple intelligences" theory and the proficiency level in order to Dreyfus's Proficiency stages to address each learner's unique learning desires and requirements [50]. Shih et al. [50] explores how an e-learning system has a personalized interface, enabling different student views/learning objects and activities by applying Gardner's "Multiple intelligences" theory, Dreyfus's Proficiency stages focusing on a heterogeneous student group.

As well as applying adaptivity to user interfaces is also a method to increase students' engagement and to provide better educational outcomes further enhancing personalization. Recent studies have discovered different approaches to add adaptivity to user interfaces including usage of AI, User Modelling(UM) and using HCI for multimodal interfaces [4]. Intersection of personalization and adaptivity improves the usability and effectiveness of user centered learning. LMSs play a crucial role in assisting personalized education. Among the many technological advancements, Adaptive User Interfaces(AUI) has become a key part in keeping students interested in their studies. Bagustari and Santoso [4] is a study which analyzes the development of AUI in education 4.0 for improving LMS to facilitate the dynamic changes in education following the approaches of AI, UM, HCI and using multimodal interfaces applied to create AUI. Further, Bagustari and Santoso [4] discusses application of technological and heutagogical(individual) aspects in their evaluations.

Many different studies have been conducted in different areas focusing on leveraging personalized education using various technological advancements. Alves et al. [1] discusses how "personality" is important as a differentiating factor in user interface designing. According to Alves et al. [1], personality is defined as the collection of habitual behaviors, cognitive and emotional patterns that result from biological and environmental influences and it is important for fulfilling diverse user needs in customizing user interfaces. It also discusses the application of user personality, focusing on improving Graphical User Interface(GUI) design exploring the application of hybrid mental models and introvert, extravert, neutral GUI.

When designing user interfaces it is important to explore and follow best practices and theories to provide better user experience. It is crucial to follow Nielsen's usability heuristics, user centered designing principles, and color theories [34, 21]. There are different learning platforms that have been discovered to enhance personalized learning. The development of personalized learning platforms has accelerated in recent years due to the rise of artificial intelligence (AI). Machine learning algorithms are now widely used in educational tools to dynamically modify learning pathways in accordance with user input, evaluations, and behavioural patterns. However, these platforms primarily address the needs of neurotypical learners, and there is a lack of adequate systems addressing the special cognitive and attentional needs of individuals with neurodevelopmental disorders such as ADHD.

3.2 Personalized Learning for ADHD Students

When discovering personalized learning approaches for ADHD individuals it is crucial to consider their key symptoms including impulsivity, hyperactivity, and inattention. Due to unique learning difficulties with attention issues, they are facing special challenges in conventional educational environments. Personalized learning facilitates the specific learning requirements of ADHD students focusing on their attention span and impulse control. Different studies have conducted focusing on exploring enhanced learning experiences for the ADHD students with structured, distraction-free, engaging, and manageable learning environments.

Gkora [18] highlights the importance of integrating autonomy-supportive practices and technological interventions in primary education for students with ADHD, emphasizing the need for interdisciplinary approaches. A personalized AI-based system which analyzes contextual data, identifies user behavior patterns, and effectively visualizes the information to provide insights to ADHD students, has been explored through the study Ravishankar [47]. As stated in Ravishankar [47], the project has used an iterative Double Diamond framework for incorporating co-design techniques to explore the real life problems of ADHD and ensure that the concept is tailored to the user's needs.

Numerous studies have been conducted on how general users comprehend data visualizations. Tran et al. [56] emphasizes how individuals with ADHD interpret data visualizations and how their comprehension may differ from the general users, by conducting a crowd-sourced survey involving 70 participants with and without ADHD. Through that they have found that participants' accuracy and response times are impacted by specific chart elements and ADHD. This study reveals that different chart embellishment types impact accuracy and response times for individuals with ADHD differently depending on the types of questions, suggesting visual design recommendations make accessible data visualizations for people with ADHD.

The emergence of personalized learning approaches have become helpful for the ADHD students for improving their focus and engagement levels and increasing their self-esteem and confidence. As well as it leads to enhancing their academic performance and reduced dropout rates. Various research and reviews have been conducted in this area for applying personalization for the educational environments of ADHD individuals, emphasizing selected audiences with different age groups. Most of the existing studies have been focused on leveraging personalized education of primary ADHD students, leaving adults with ADHD significantly underrepresented. Therefore through our research, we are focusing especially on adolescents with ADHD addressing their specific learning requirements.

3.3 Existing Learning Applications for ADHD Students

According to the vast technological development, different studies have been conducted and different learning applications have been implemented for ADHD students using various technologies.

3.3.1 The ADHD Augmented Learning Environment (AHA) Project

The AHA project, supported by the European Commission, intends to improve education for ADHD children by utilizing Augmented Reality and web technologies[21]. The project's goal is to improve reading and spelling skills while also increasing academic engagement by using augmented reality into online literacy programs.

AHA, which targets primary school pupils in Ireland, aims to help ADHD students stay focused, make fewer mistakes, and complete homework more efficiently. It builds on established methodologies and works with the Web Health Application for ADHD Monitoring (WHAAM) project to ensure user privacy and data security[32]. AHA seeks to provide an effective and innovative learning environment for ADHD children. It includes the key objectives of providing digital resources to teachers and parents, providing best-practice examples and integrated solutions, collaborating with associations on pilot research, creating standards and a road map for educators and legislators.

3.3.2 aTenDerAH: A Videogame Application to Support e-Learning Students with ADHD

aTenDerAH is a videogame that supports e-learning for young adults, particularly those with ADHD[30]. aTenDerAH was created with Unity, Cinema 4D for 3D modeling and animation, and Photoshop for texture creation. The video game was integrated into the Atutor e-learning platform to conduct a case study on the perception of aTenDerAH by students with ADHD, non-ADHD students, and teachers. Participants have expressed satisfaction with the tool's goals and the videogame's good impact on their learning process

3.3.3 The Facilitating On-going Concentration in Undergraduate Students (FOCUS) Device

As stated in Thomas and Cezeaux [54], FOCUS application is a prototype designed to enhance the academic performance of students with ADHD. The major objective of the device is to offer reminders that are not stigmatizing in order to assist students in maintaining their focus throughout class and while completing their studies.

As the key features of this application, the device incorporates the Texas Instruments eZ430 Chronos Development Tool, a digital timepiece that manages peripheral devices situated on the student's desk. These peripheral devices consist of vibration motors, LEDs, and speakers that are capable of emitting reminder signals in the form of vibrations, beeps, and flashing lights. As well as, it was specifically designed to be compact, lightweight (weighing less than two pounds), and inconspicuous, based on a survey done among students with ADHD. The reminders are randomized to thwart the student's habituation and disregard of the signals. This device is engineered to have a height and thickness of less than five inches, and it has a battery life that may last anywhere from one week to one month, depending on usage. The FOCUS application is designed to offer a convenient and efficient solution for students with ADHD, assisting them in effectively regulating their focus and enhancing their academic performance, all while avoiding the negative connotations sometimes associated with traditional education treatments.

3.3.4 X5Learn: A Personalised Learning Companion at the Intersection of AI and HCI

X5Learn is a human-centered educational platform which is implemented by integrating AI and HCI supports browsing and accessing online educational resources [42]. It is available at 'https://x5learn.org X5Learn' is an educational tool designed to support both teachers and students. It offers tools for interacting with open educational videos and pedagogical preferences. Teachers can reuse, revise, and redistribute courseware, such as videos, PDFs, and exercises. Students can select content, make notes, and write reviews, and use a scaffolded interface. The tool also features a personalized recommendation system to optimize learning paths and adapt to users' preferences.

It has been designed as a suite of interconnected tool including,

- X5GON Connect Service: This connect service is used to connect OER (Open Educational Resources) repositories globally into a single network, enabling users to search and browse all OERs.
- X5GON Translate: Using artificial intelligence (AI) techniques, this tool automatically creates transcriptions and translations in various languages through automatic speech recognition.
- X5Learn: This is the main learning platform where the AI search engine and

recommendation system return educational resources. Thousands of video lectures are available in this database.

- X5GON Discovery: This is an AI-powered search engine that allows users to quickly find educational resources by entering important parameters.
- X5GON TrueLearn: An AI-powered recommendation engine that offers clear educational suggestions for lifelong learners by utilizing TrueLearn

3.3.5 Therapy Tool for Adolescents with ADHD

Children and adolescents who have ADHD must receive treatment because, if left untreated, symptoms may last into adulthood. This study Nugawela et al. [40] explores and has developed an Android app to monitor patient progress and offer therapeutic treatment for ADHD symptoms. As stated in this study, one person participates in each testing module to ensure accuracy and remove subjectivity. The app addresses all of the primary symptoms of ADHD and provides a workable way for patients, parents, and physicians to track their progress. In addition to offering a therapeutic treatment for ADHD, the app keeps track of the performance and development of its users. This solution is helpful for preventing the progression of ADHD symptoms into adulthood.

3.3.6 ECHO ADHD in India: A Feasible and Acceptable Training Model for Child-care Physicians to Identify and Manage Attention Deficit Hyperactivity Disorder

A case based tele-mentoring model which is designed to address gaps in physician knowledge and skills regarding ADHD has been suggested through ECHO(Extension of community health outcomes) ADHD [46]. The ECHO ADHD model has become the result of the advancements of the initial ECHO model, which has been gradually developed through several studies and testings. As the initial step, University of New Mexico Health Sciences center developed Project ECHO to enhance the outcomes of the individuals with Hepatitis C virus infection in underserved areas. There, the participants were connected via secure video conferencing to a specialist team. After succeeding this initial project, the ECHO model was used globally and in India to train physicians to manage a range of complex medical and behavioral conditions. Then the ECHO model was culturally adapted in India to enhance physician capacity for diagnosing and managing developmental conditions like autism spectrum disorder, but no similar program has been developed for ADHD in India at that time.

In this study, the authors have created a pilot ECHO ADHD program in India, incorporating key learnings from ECHO Autism India. The program included parent and self-advocates in the hub team and emphasized a strength-based, family-centered approach during didactic and case-based discussions. ECHO ADHD pilot round was initiated in the midst of a COVID-19 infection in India in 2021. Used high-quality secure video conferencing technology to connect participants.

The purpose of this study was to assess the ECHO model's acceptability and feasibility for training participants in evidence-based practices (EBP) for ADHD. At the beginning and end of the program, it also evaluated views regarding the appropriateness and viability of the role in diagnosing and treating ADHD, as well as changes in self-efficacy in recognizing and managing ADHD and comorbidities.

3.4 Use of Generative AI (GenAI) in Personalized Learning

Artificial Intelligence (AI) has become a great potential for enabling more personalization in the educational sector. Machine Learning(ML) techniques, GenAI, LLM which are subsets of AI are widely used in leveraging personalized education platforms. With the development of dynamically adaptive educational experiences, GenAI has recently started to transform personalized learning. There are different model architectures including Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), Diffusion models. Different GenAI tools are commonly used in reshaping educational platforms such as Text generation tools including GPT, Jasper, AI-Writer, Image generation tools including Dall-E 2, Midjourney and Stable Diffusion.

Different studies have been carried out and discovered usage of GenAI in personalized education. Yu et al. [62] paper discusses how AI advancements can improve MOOC learning and research. It highlights that AI-enhanced Massive Open Online Courses (MOOCs) allow students to have a more personalized learning experience with personalized learning paths, activities and allowing them to access learning content without pre-defined sequences. AI-driven personalized learning systems have been explored integrating intelligent agents, automatic scoring, assessment and learner-support chatbots. In the study Perez-Ortiz et al. [42] , a human-centered AI powered platform is explored for supporting access to free online educational resources. The systematic literature review Castro et al. [11] examines the key drivers of personalized learning, a pedagogical approach adapted to individual needs, and the role of AI in reinforcing these drivers.

The study Pesovski et al. [43] emphasizes personalizing learning materials in education. They proposes a tool within a software engineering college's LMS that generates learning materials based on the learning outcomes provided by professors for specific classes, with the usage of LLMs using OpenAI's API. Logacheva et al. [28] discusses a user study in an introductory programming course that evaluated the quality of contextually personalized exercises created with GPT-4, revealing high student engagement and attitudes towards the system, indicating its effectiveness. The integration of LLMs such as ChatGPT with robotic assistance in intelligent environments to improve ADHD nonpharmacological therapies is explored through the research Berrezueta-Guzman et al. [7], with an emphasis on theory, practical application, and ethical considerations.

Barua et al. [5] evaluates AI-assisted tools for addressing learning challenges in students with Neurodevelopmental Disorders(NDD). It highlights their effectiveness in improving social interaction and supportive education, and recommends future tools for personalized learning. They have used conventional AI methods including Deep learning, ML and Advanced AI methods including deep models - CNN,LSTM. Sukiman and Aziz [52] discusses the integration of AI features in personalized learning interventions for students with learning difficulties, their implementation, and the additional features required for successful delivery.

3.5 Use of GenAI in Interface Generation

Recently with the advancements of Generative AI, the use of AI models for UI generation have significantly improved. The study Goloujeh et al. [19] explores the prompt journey of text-to-image AI tools, focusing on user experiences. Through semi-structured interviews with 19 users, the study provides insights into writing, evaluating, and refining prompts. Rajcic et al. [45] discusses the role of prompt-based interfaces in artists' creative practice, focusing on seven visual artists who used AI for ideation and production, and presents ethical design considerations for future GenAI interface development.

Through the study Liu and Chilton [27], conducts five experiments on prompt engineering for text-to-image generative models, revealing significant differences in generation quality and failure modes, and identifying hyperparameters with significant effects. Weisz et al. [59] presents six principles for designing GenAI applications, addressing unique UX characteristics and extending existing AI design issues. As stated in Weisz et al. [59], they have been derived according to "intent based outcome specification" by Nielsen [35]. We apply these design guidelines and principles in our work to generate advanced user interfaces to provide optimum learning experiences for ADHD students.

Chapter 4

Methodology

The research was conducted in 6 main phases, aligned with the DSRM [41]. The Design Science Research (DSR) approach focuses on expanding knowledge through the creation and evaluation of innovative solutions to address particular, real-world problems [57]. Designing practical artifacts and developing design theories are key components of the DSR approach, which makes it relevant for IS research [6]. According to vom Brocke et al. [57] DSR's, the output consists of both the newly created artifacts and design knowledge, which offers a more thorough comprehension of the reasons behind the artifacts' enhancement (or disruption) of the relevant application contexts through design theories. The DSR method includes an iterative process with six steps,starting from the problem identification and motivation of the research to the final evaluation and communication [41].

The DSR approach was chosen as the methodology for our research because it is an IS study that aims to achieve both scientific rigour and practical relevance. By applying this approach, our study aimed to explore the best personalized learning interfaces tailored to the unique needs of ADHD students, leveraging the potential of generative AI. The DSR approach offers a pragmatic methodology that is flexible, focused on outcomes, and able to handle the complexity of the real world. It allows integration of diverse methods to address research questions based on their suitability and context, making it ideal for designing and evaluating multiple learning interfaces tailored to the needs of individuals with ADHD [24]. In this study, we demonstrate and evaluate the methodology by presenting and comparatively analyzing three distinct experiments: a conventional LMS used by the general population, a manually optimized LMS designed specifically for ADHD users, and an AI-driven LMS customized through enhanced prompting techniques. By utilizing DSR, this study ensures that each interface is rigorously designed, developed, and evaluated, aligning with both technological breakthroughs and scientific evolution. By leveraging DSR for our research, we were able to actively engage in iterative design and development cycles, informed by HCI principles, design guidelines and generative AI techniques. Through the use of this DSR methodology, this study aims to explore the most effective learning interfaces for enhancing the learning experience of individuals with ADHD.

4.1 Problem Identification and Motivation

The initial step of our research involved problem identification and motivation which was focused on evaluation of needs and requirements gathering. A comprehensive literature review was undertaken to study the biological background and unique needs of ADHD individuals, and the current state of the traditional educational environments. According to the findings we identified a notable absence of learning platforms for addressing the special learning abilities of ADHD individuals through exploring personalized learning interfaces. It was discovered that there is a significant lack of research in the area of addressing the specific learning requirements of adult ADHD individuals, although some of the research has been conducted focusing on the primary education of ADHD students. It emphasised the significant need for the exploration of a personalized learning environment for ADHD students, tailored for addressing their specific learning requirements to enhance their overall educational outcomes. Further, we discovered that although generative AI shows immense potential, still now its application for creating ADHD-friendly LMS interfaces remains largely unexplored, especially not focusing on generating personalized learning interfaces.

We conducted a survey and identified ADHD individuals within the age range of 18-30, including medically diagnosed and self-diagnosed participants and studied the educational difficulties they face in their existing learning environments.Through the survey and interviews, we got an understanding of the unique learning requirements and obstacles faced by ADHD individuals and we determined the essential functions and features required for designing the personalized learning interfaces for them. The insights gathered from these resources further underscored the significance of exploring an optimised learning solution revolutionising the potential of generative AI for the adult ADHD students to excel in their academics.

4.2 Defining Solution Objectives

The lack of adaptive learning platforms and tailored learning interfaces for students with ADHD highlights a significant gap in the current conventional educational environments. The particular cognitive and behavioural difficulties that ADHD learners encounter, such as impulsivity, hyperactivity, and inattention, are frequently ignored by traditional LMS. This underscores the pressing need for comprehensive and accessible interventions that can provide equitable learning opportunities for them to excel in their academics.

In addressing this critical gap, our research aimed to bridge the divide and make a substantial impact on society. Our primary goal was to explore the potential of generative AI to revolutionize personalized learning interfaces for individuals with ADHD. Through that we expect to create an innovative, adaptive learning environment that meets the special needs of students with ADHD, improving their motivation, engagement, and overall academic performance.

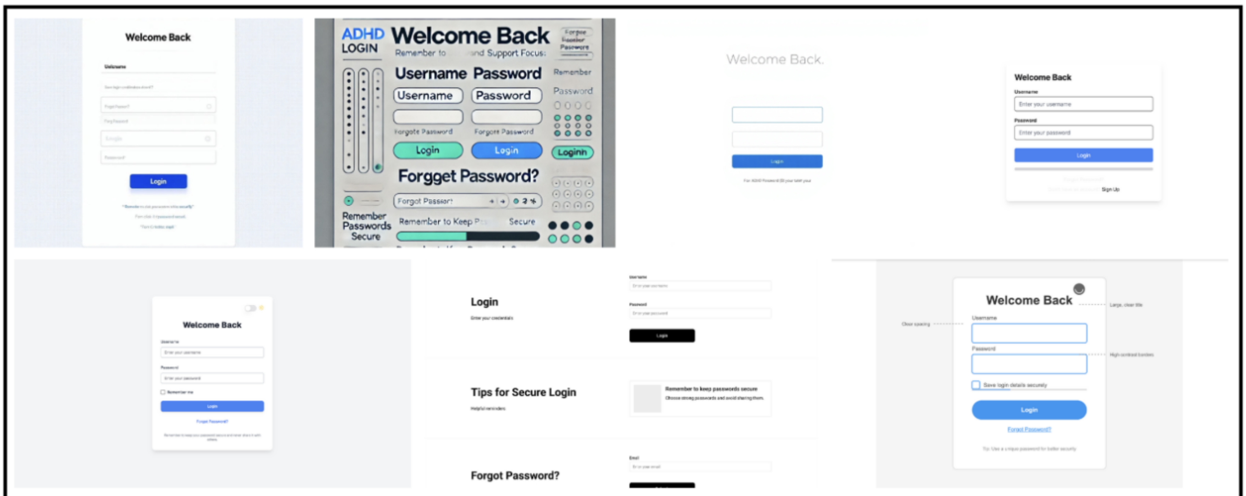
In this step, we identified the ideal GenAI tools and techniques that can be utilized to fulfil the specified requirements of ADHD students. We evaluated several GenAI technologies according to factors including accuracy, adaptability, user-friendliness, and ease of integration.

4.2.1 Explored AI Models and Tools for UI Generation

When choosing an AI model for UI generation , we explored several mostly used AI models, such as DALL·E, Flux AI, Vercel VO, Open UI with GPT-4.0, Figma, and Claude AI, along with the Stable Diffusion 3.5 model from StabilityAI in Hugging Face space.

We used 3 different prompts to identify the models behaviour in generating these user interfaces and how the model responds to various prompts. With the results we send out a survey including the prompt and response as an option to experts and non-experts in UI/UX design. We got 25 total responses and with those responses, highest scored models were Vercel VO and Claude AI. With those results, we decided to conduct our research further with 3 models combined with Claude AI, Vercel VO and ChatGPT 4.0. Main parameters we tested with user interfaces were accuracy, detail, artistic, relevance and creativity with of the generated interfaces with respect to the prompt given. The prompts were designed according to the design guidelines presented by McKnight [31].

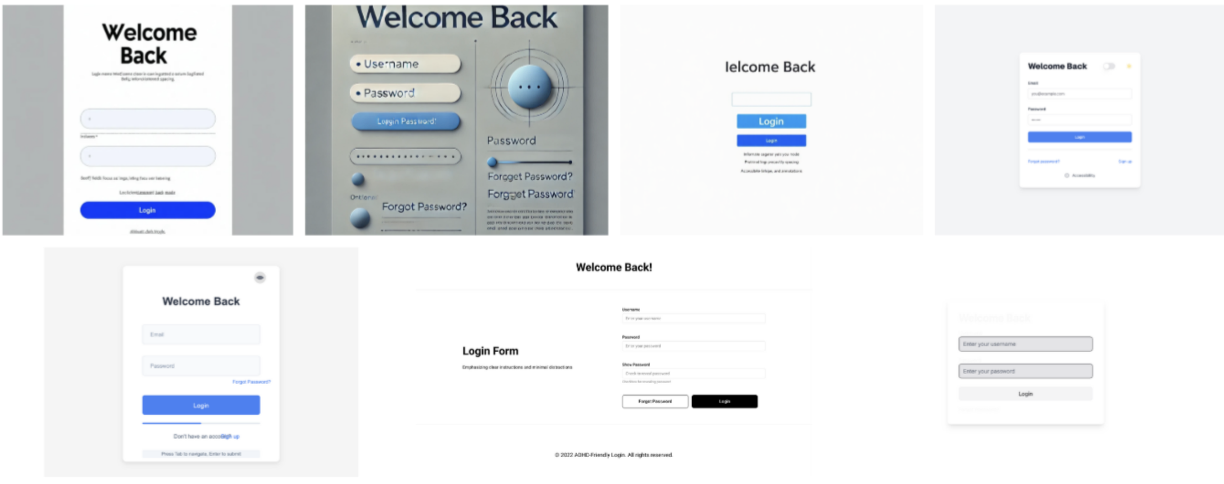
Test 01 Prompt:
"Generate a wireframe of a login page designed for users with ADHD, featuring a clean, minimal layout that prioritizes simplicity and focus". See figure 4.1.



a) Results for prompt 01

Figure 4.1: Prompt 01 Interfaces

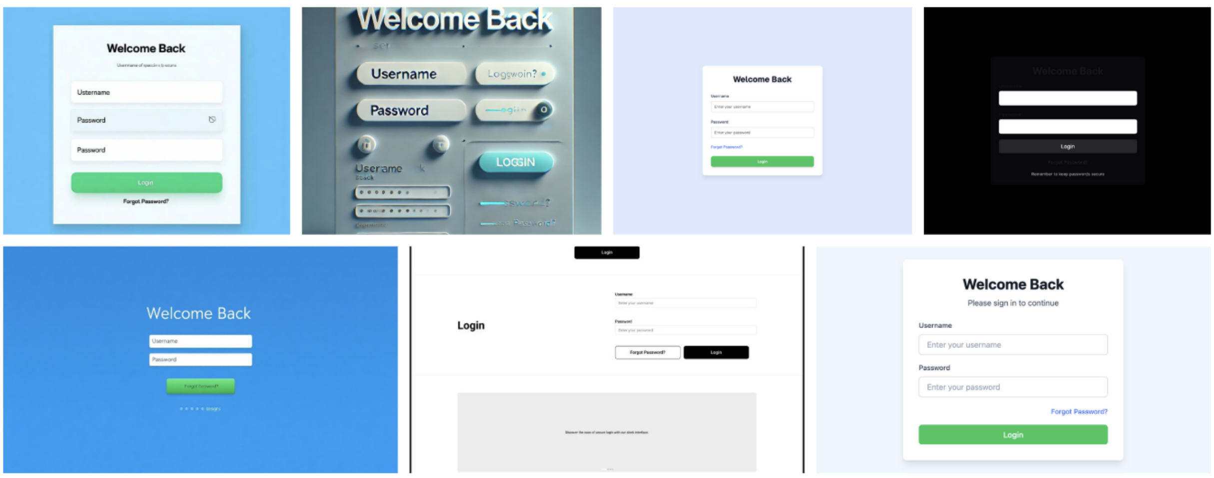
Test 02 Prompt:
"Create a minimalistic wireframe of a login page with a calming colour scheme, featuring a bold "Welcome Back" header, contrasting input fields, and a soothing blue 'Login' button. The background is soft grey, ensuring clear, uncluttered spacing, and includes an optional dark mode toggle. An informative progress bar is subtly integrated beneath the button, while accessibility links, tips, and annotations enhance user focus and clarity, all outlined in a clean, modern design". See figure 4.2.



b) Results for prompt 02

Figure 4.2: Prompt 02 Interfaces

Test 03 Prompt:
"A sleek, minimalistic login page with a calming blue background, featuring a bold "Welcome Back" header. Two high-contrast form fields for 'Username' and 'Password' sit neatly above a soothing green 'Login' button, flanked by a subtle 'Forgot Password?' link. Soft shadows enhance spacing, while a discreet loading bar awaits user engagement, promoting focus and clarity in design". See figure 4.3.



c) Results for prompt 03

Figure 4.3: Prompt 03 Interfaces

Results of AI-Generated Interface Evaluation Survey

We conducted a survey in choosing the best AI model to use in user interface generation. We tested 7 AI models as shown below in Table 4.1 with 25 users in total which consisted of UI/UX professionals and non experts. The best-performing model according to the survey result was the Vercel VO AI model as indicated by Figures 4.4,4.5,4.6. Given that both Claude AI and ChatGPT demonstrated favourable outcomes, we selected to prioritise these three models in generating user interfaces. By combining multiple models, we managed to leverage the unique capabilities of each model, such as Vercel VO for creativity, Claude for long-form reasoning and ChatGPT for structured responses. The approach helped us in reducing the risk of model-specific errors and blind spots.

Image	**AI Tool**
Image 1	DALL·E
Image 2	Flux AI
Image 3	Open UI with GPT-4.0
Image 4	Figma
Image 5	Vercel Vo
Image 6	Claude AI
Image 7	Stable Diffusion

Table 4.1: Model Evaluation Results

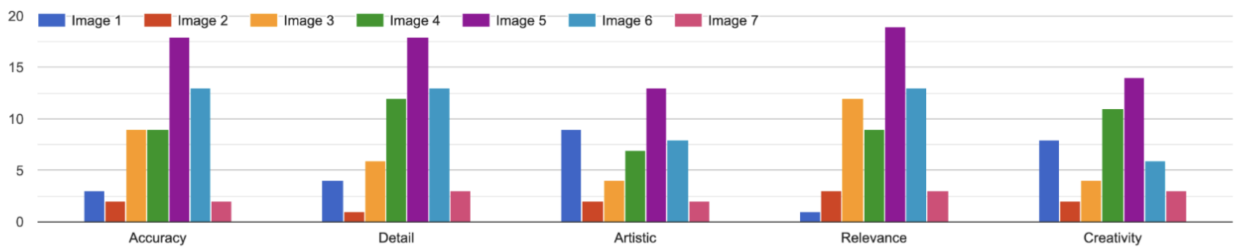


Figure 4.4: Test 01 Results

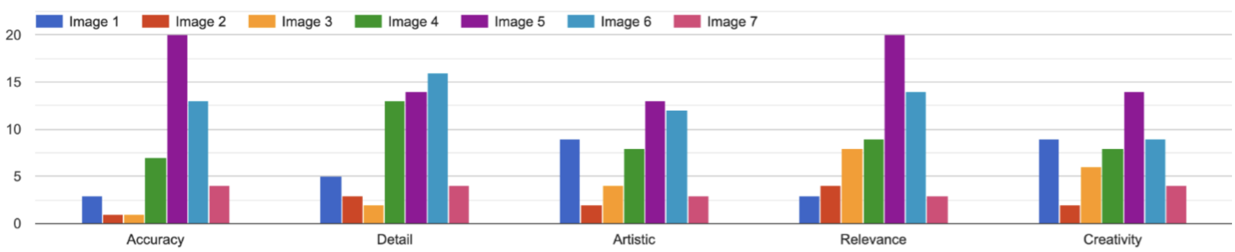


Figure 4.5: Test 02 Results

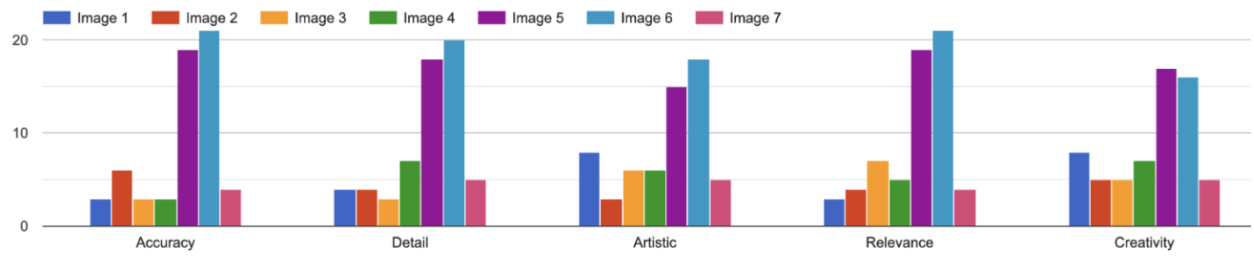


Figure 4.6: Test 03 Results

4.3 Design and Development

This step of the research focused on the design and development of the prototypes. We designed and developed three distinct prototypes: a conventional LMS used by the general population, a manually optimised LMS designed specifically for ADHD users, and an AI-generated LMS customised through enhanced prompting techniques. When designing each of the prototypes we applied Nielsen's Usability Heuristics[36, 33], HCI theories, UI/UX best practices and ADHD-friendly design guidelines. We followed an iterative development approach for the development of the prototypes, incorporating constant feedback from specialists and potential users.

For the design and development of the three LMS prototypes, the subject "Philosophy of Science" was selected as the core content area. This subject was intentionally chosen because it is relatively uncommon and unfamiliar to most students, thereby minimizing prior knowledge and reducing potential bias in the results. By using a subject that students had not previously studied, the analysis of interaction, engagement, and quiz scores more accurately reflects the effectiveness of each prototype's design, rather than differences in familiarity and existing knowledge of the subject. By doing this, we ensured that the comparative evaluation focused on the effectiveness and educational impact of the LMS prototypes themselves.

4.3.1 Design Guidelines and Features of ADHD Friendly Design

According to the findings of the comprehensive literature review we found ADHD friendly features and design principles derived through different studies. McKnight [31] proposes a series of 15 guidelines for ADHD friendly designs, which are aligning with other usability guidelines. Neuro-Inclusive UX Design Principles have been suggested in Team [53], explaining the difference between neurodiverse and neurotypical users. With the use of MCKnight's design guidelines, Phalke and Shrivastava [44] proposes a categorization of visual design guidelines that can be applied in Health Informatics (HI) applications. We have applied these principles, ADHD friendly design features in designing our learning interfaces, tailoring them to the specific learning abilities of the ADHD users.

4.3.2 Usability Heuristics and HCI Principles

We have explored various HCI standards and UI/UX best practices during the literature review and each of the prototype UI designs was refined with them to ensure optimal usability and accessibility for students with ADHD. Throughout the design and implementation phases of our study we have followed, 10 usability heuristics devised by Jakob Nielsen [36],[33]. when designing the user interfaces. Further we followed error message guidelines [37], accessibility guidelines [20], visual design principles and user-friendly color schemes according to the findings of our literature review and the insights taken from UI/UX experts.

4.3.3 Prompting Techniques

When using prompting techniques, we used both zero-shot prompting and few-shot prompting while generating content. In the initial interface generation, zero-shot prompting was used without providing any examples of layouts and giving direct requests for enhanced features. This was combined with a high-level abstraction prompting technique where we would ask for a conceptual layout instead of exact UI code. The focus was on goals and user personas.
E.g : "Make the profile page more accessible, with editable fields, full width, an editable profile pic, appropriate colors, ability to save changes."

To improve the functional and visual requirements of the interfaces, we used constraint-based prompting.
E.g : "specifying functionalities such as "editable fields", "fill the page", "appropriate colours", "editable profile pic", and "save changes."

Accessibility prompting technique was followed while giving prompts. We observed that including keywords such as "ADHD friendly" and "accessible" results in neurodivergent-friendly concepts embedded in results such as colour contrast, focus states and clarity. While following persona-based prompting, we used the keywords "ADHD users" and "ADHD students" to reference that the user interfaces are focused on their needs.

After generating initial layouts, mostly the task chaining technique (multi-step prompting) was used in order to describe interaction sequences in user interfaces and generate dynamic behaviour. Eg : button → popup → update → feedback. In some instances, behavioural context prompting techniques such as specifying user experience outcomes from an interaction, such as "give success message", were used to guide the AI.

Prompting template:
'Design an [ADHD-Friendly] [Component] with [Accessibility feature] and [Interaction flow].'

4.3.4 Prototype Development – Prototype 01: Base LMS

The first prototype developed in this research was a Base LMS, which served as a control version representing conventional learning management system typically used by the general student population. This prototype was built using React for dynamic component rendering and Tailwind CSS for utility-first, responsive UI design. The Base LMS contained essential features such as user login, course listing, calendar view, and assignment tracking. However, no ADHD-specific design considerations or accessibility improvements were incorporated at this stage.

The purpose of creating the Base LMS was to establish a benchmark for comparing how traditional learning environments cater to students with neurodivergent learning patterns. The layout, navigation structure, and visual aesthetics followed standard UI/UX conventions without applying ADHD-centered heuristics. Color schemes were neutral and layout spacing, element focus, and interaction feedback mechanisms were minimal, reflecting a typical LMS environment. This prototype served as the baseline model for comparing usability and effectiveness against the ADHD-optimized versions.

While functional and structurally complete, this version lacked personalization, feedback mechanisms, and interface elements optimized for cognitive load management or impulsivity reduction which are key factors for ADHD users. This prototype was subjected to user testing among both neurotypical and ADHD-identified individuals to assess usability gaps and highlight improvement areas. This Base LMS acted as the control prototype to highlight the limitations of traditional LMS designs for ADHD learners and provide a comparative foundation for evaluating improvements in the manually and AI-enhanced prototypes. The insights gathered from this testing phase informed the development of the second and third prototypes.

4.3.5 Prototype Development - Prototype 02: Manually Improved LMS

The second prototype was a manually improved LMS, tailored specifically for students with ADHD to address ADHD-specific challenges. This version was also developed using React and Tailwind CSS, but it incorporated ADHD-friendly design principles, informed by the literature review, survey findings, and expert consultations in HCI and accessibility. This prototype introduced targeted optimizations for enhanced engagement, attention retention adn reduced cognitive load.
This version applied several design adaptations including:

- Reduced visual clutter and simplified navigation flows
- Micro-interactions and animated feedback to maintain engagement
- Time management and task prioritization widgets

- Enhanced typography, progress visibility, and enhanced feedback systems
- Consistent spacing and layout structuring to minimize cognitive overload

We used Nielsen's Usability Heuristics, HCI principles, Web Content Accessibility Guidelines (WCAG) 2.1 standards[60] and McKnight's ADHD Design Guidelines[31] to revise component layouts and interaction flows. For example, the course content page was restructured to limit distractions and emphasize important sections, while each and every page included features like prioritization toggles and visual indicators for overdue tasks.

Moreover, we incorporated design elements inspired by neuro-inclusive UX standards, expert evaluations and user feedback such as visual focus indicators, predictable user flows, and error-tolerant forms. The prototype was manually iterated through several design review cycles with feedback from ADHD participants, ensuring its alignment with the lived experiences and preferences of neurodivergent learners.

This prototype bridged the gap between generic LMS platforms and highly personalized learning interfaces by offering a user-friendly, inclusive digital environment. It was critically evaluated alongside the Base LMS and AI-enhanced LMS(prototype 03) to measure the degree of changes in usability, effectiveness, accessibility, and user satisfaction among ADHD students.

4.3.6 Prototype Development – Prototype 03: AI-Generated LMS

The third prototype used LLMs to automatically generate neurodiversity-aware LMS interfaces, with a focus on Nielsen's usability heuristics, HCI principles, and McKnight's ADHD design guidelines and compliance with WCAG 2.1.To achieve this, we mainly created a ChatGPT-4 extension programmed with domain-specific constraints to generate low-distraction UI components through iterative prompt engineering and used Claude AI and Vercel VO, parallely addressing the complex interplay between universal accessibility standards and neurocognitive-specific design requirements.

Architectural Framework

The main technique used for the user interface generation with LLMs is UI code generation, which is also known as UI layout generation. This terminology is prevalent in recent research, such as the UICoder framework, which specifically focuses on fine-tuning LLMs to generate user interface code through automated feedback mechanisms [61].

With our system, the prompts enforced keyword triggers (”students”, ”ADHD”) to maintain context awarenes, generating features as listed below,

- High-contrast navigation systems using #0052CC blue for active states against #FFFFFF white backgrounds
- Fixed-position icon grids with 12px spacing consistency
- Progressive disclosure mechanisms for complex workflow steps
- Color theme switching for users with cognitive sensitivity
- Resizable Text enabling supports users with low vision or reading difficulties.
- Page Zoom feature
- Interactive Learning Modules/content presentation
- Accessibility Options Panel That centralizes all accessibility controls, allowing users to customize their experience according to individual need.

Interactive refinement process of UI generation

The initial outputs generated by LLMs were not up to web application standards because the LLMs often struggle to generate systems as a whole and generate the systems page by page. With that, the credibility of a system decreases and the error margin increases [58].
Initial AI outputs exhibited three critical limitations such as,

- Lack of understanding on system architecture of a web application [10].
- Over-reliance on monochromatic schemes (68% white/blue combinations)
- Inconsistent API endpoint mapping across React 18 components.

The development team implemented a feedback loop where,

- Generated interfaces underwent automated axe-core accessibility testing
- Failure points were translated into reinforced prompt constraints
- Claude AI and ChatGPT re-generated components with error-specific corrections

Notably, 80% of successful prompts required explicit "ADHD Friendly" modifiers to override default Material Design tendencies toward complex visual hierarchies.

Generative UI Limitations

The prototype revealed fundamental challenges in AI-driven accessibility design,

- Context blindness where Models failed to propose focus management tools like "Attention Zones" without explicit examples

- Asset generation limitations, which included images, required manual prompt injection, as LLMs performed relatively low with image inclusion.
- Contrast paradox with WCAG-compliant color pairs often contradicted ADHD-friendly, low-arousal palettes

4.4 Demonstration

The demonstration phase of the research unfolded iteratively, This research's demonstration phase was carried out iteratively, ensuring that the developed prototypes were rigorously tested, improved, and verified in real-world environments. This phase involved feedback collection, and refinement in order to ensure that the learning interfaces satisfied the particular requirements of people with ADHD while maintaining their usability and effectiveness levels.

We shared our three distinct prototypes: the conventional LMS, the manually optimized LMS, and the AI-generated LMS with our ADHD user group of 22 participants, and gathered their feedback on each of the learning experiences. Engaging the target audience in the evaluation process allowed for the acquisition of important insights regarding the usefulness and impact of each prototype in actual educational environments for exploring the optimal learning environment for the ADHD students. Based on this feedback insights, evaluation was done using a variety of methodologies.

All the 22 participants with ADHD was first given the Base LMS to ensure a common starting point and for unbiased baseline data collection. To reduce order effects and minimize the bias in evaluation of the other 2 prototypes, the counterbalancing approach was used to control learning and fatigue effects that could skew the results. Half of the user group received the Manual LMS first, while the other half received the AI LMS first. After each prototype session, the users were given the SUS, UEQ and Quiz score forms to gather quantitative data. After the completion of all the 3 prototypes a final interview was conducted with each participant to gather qualitative insights. The combination of gathering of all these data provided a holistic view of each prototype's usability and effectiveness.

4.5 Evaluation

Under this phase DSR methodology, the research commenced with the evaluation and testing stage. Here we analyzed the data collected from the demonstration phase, and determined the prototype's usability and effectiveness based on the derived insights. We measured usability, accessibility, and other needed metrics. The effectiveness of the personalized learning interfaces was assessed by comparing pre-test and post-test results to evaluate improvements in usability, effectiveness, engagement, and learning outcomes.

4.5.1 Ethical Considerations

Ethical considerations were paramount throughout the evaluation process. Informed consent was obtained from participants, ensuring transparency about the study's purpose and procedures. The privacy and confidentiality of participants were protected, with anonymized data used for analysis and reporting. The study adhered to the ethical guidelines set by the University of Colombo School of Computing Institutional Review Board (UCSCIRB), and any potential risks or discomfort to participants were minimized.

The research team remained committed to conducting the evaluation ethically and responsibly, prioritizing the well-being and rights of the ADHD students involved in the study.

4.6 Communication

In this stage, the primary objective is to effectively disseminate the valuable insights and knowledge obtained from this study through curated communication channels to relevant stakeholders. We expect to share our findings through professional networks, presenting at conferences, thesis writing and research publications, contributing to the field of educational technology and ADHD research. We are currently working on publishing our research in academic journals, and this forthcoming publication will focus on evaluating the effectiveness of personalized learning interfaces in enhancing the learning experience of ADHD students.

Chapter 5

Results and Evaluation

The evaluation phase of this study involved systematically analyzing the usability and effectiveness of the proposed LMS prototypes, by collecting user feedback and evaluating through both qualitative and quantitative metrics. This study was conducted in accordance with the ethical research guidelines provided by the Ethics Review Committee (ERC) of the University of Colombo School of Computing and with ERC approval.

5.1 Participants

The study was conducted with a diverse group of ADHD students within the age range 18-30, who have prior experience with online learning platforms. We selected these students from different educational backgrounds to avoid biases of the evaluations. At the beginning of the research, we conducted a survey to identify the needs ADHD students, and based on gathered responses, we created a user group of 26 ADHD students consisting of 7 medically diagnosed ADHD students and 19 self-diagnosed ADHD students using the ADHD test tool, which has been introduced by Mental Health America (MHA) [2].

According to the preliminary survey, a total of 93 students from various academic backgrounds participated. Among them, 43 individuals (11 medically diagnosed and 32 self-diagnosed) were identified as students with ADHD. Out of these, 26 students consented to participate in the evaluations of our study, forming the user group. This user group consists of 12 female students (4 medically diagnosed and 8 self-diagnosed) and 14 male students (4 medically diagnosed and 10 self-diagnosed).

We identified the severity levels of the ADHD students using the Adult ADHD Self-Report Scale (ASRS) v1.1, which was developed in collaboration with the World Health Organization (WHO)[38, 25]. Based on the participants' ADHD severity levels, individuals who fell within the high positive range were excluded. As a result, the final user group consists of 22 individuals with typical ADHD traits, as detailed below.

Participant ID	**Gender**	**ADHD Severity Score**	**ADHD Level**
AD01	Female	15	Low positive
AD02	Male	13	High negative
AD03	Female	9	Low negative
AD04	Male	15	Low positive
AD05	Male	17	Low positive
AD06	Male	16	Low positive
AD07	Female	17	Low positive
AD09	Female	15	Low positive
AD10	Male	15	Low positive
AD11	Female	16	Low positive
AD13	Male	17	Low positive
AD14	Male	16	Low positive
AD16	Male	13	High negative
AD17	Male	6	Low negative
AD18	Male	12	High negative
AD20	Male	16	Low positive
AD21	Female	10	High negative
AD22	Male	17	Low positive
AD23	Female	15	Low positive
AD24	Female	12	High negative
AD25	Female	15	Low positive
AD26	Female	16	Low positive

Table 5.1: ADHD user group

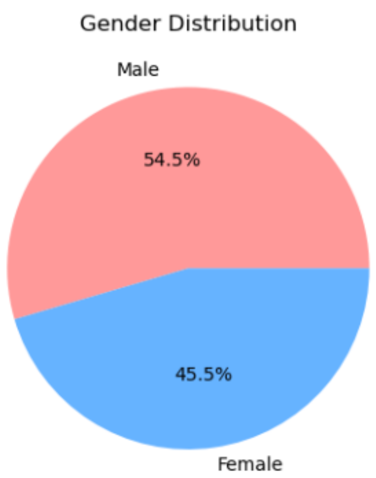


Figure 5.1: Gender Distribution

This selected ADHD user group consists of 10 female students (45.5%) and 12 male students (54.5%), showing that there is no significant gender bias in this user group. As indicated by the figure 5.1 the gender distribution of the user group is nearly balanced, with a slight majority of male participants. This balanced representation ensures that the evaluations of the three LMS prototypes are not significantly influenced by gender differences, allowing for a more objective assessment of usability and effectiveness for ADHD students.

According to the ASRS v1.1, there are four major types of ADHD severity levels, introduced in collaboration with WHO, as we stated in Table 2.2 [38, 25].

- Low negative
- High negative
- Low positive
- High positive

Based on this scale, individuals with a score of 0–17 have been identified as normal ADHD people, and those with a score of 18–24 are considered as high positive. As mentioned above, in this study we have excluded the high positive range participants to perform reliable evaluations, reducing inconsistencies and challenges.

The histogram of figure 5.2 displays the distribution of ADHD severity scores of the user group, ranging from 6 to 17. This distribution suggests our sample represents a typical ADHD population. And the bar chart shows the categorical ADHD level frequencies. The majority of participants are in the "Low positive" category, which may represent mild ADHD symptoms. The distribution across categories allows us to examine if different LMS types work better for different ADHD classifications.

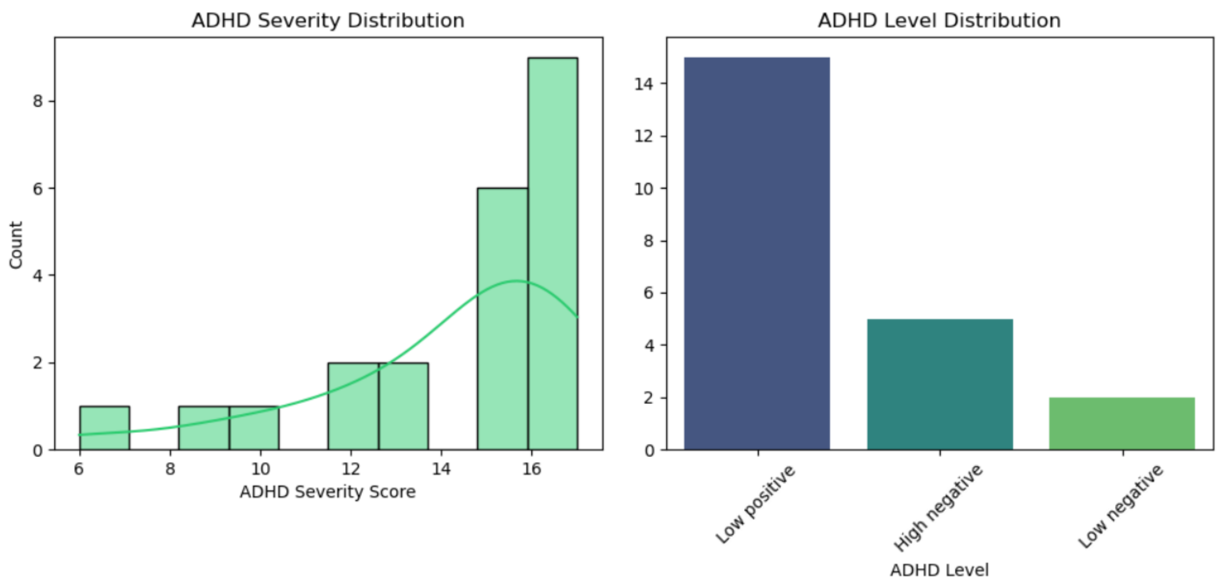


Figure 5.2: ADHD Severity Level Distribution

5.2 Usability of the System

Usability of each prototype was measured mainly by using SUS and UEQ. We gathered the SUS and UEQ evaluations performed by the 22 individuals of the ADHD user group for each of the prototypes.

5.2.1 System Usability Scale (SUS)

The System Usability Scale (SUS) is a simple ten-item scale that provides a broad perspective on the subjective evaluations of usability [9]. It is a Likert scale, which includes forced-choice questions with a statement and a 5- or 7-point rating system for the respondent's level of agreement or disagreement. According to the studies SUS has been shown as a valuable evaluation tool, due to its robustness and dependability. It indicates a good correlation with other subjective usability metrics. We have shared this SUS questionnaire for our user group and based on their scores for each of the prototypes, compared the usability applying more statistical analysis methods.

When collecting the SUS questionnaire responses from the users we have included random Attention check questions as shown in figure 5.3, to avoid the random guesses of careless answers, ensuring that the users have rated the questions after carefully reading them with attention. It helps to ensure the reliability of the responses, which leads to performing accurate statistical analysis.

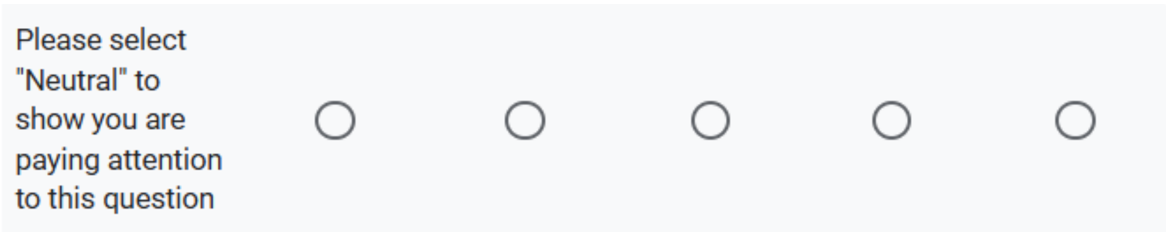


Figure 5.3: Attention check questions for SUS

5.2.2 User Experience Questionnaire (UEQ)

The primary objective of UEQ is to enable quick and immediate user experience measurement. Hedonistic and pragmatic quality factors are considered in applying UEQ [49]. It has been constructed as a user experience questionnaire containing 6 scales with 26 items.

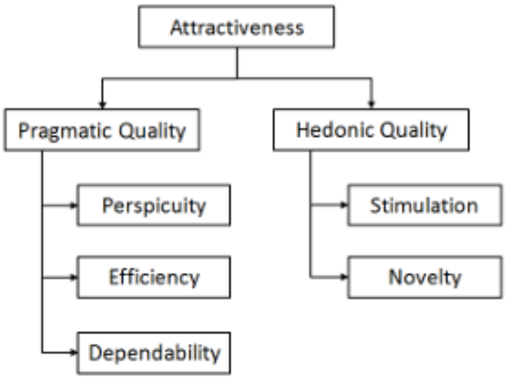


Figure 5.4: Assumed scale structure of the UEQ

In this study we have collected UEQ results from the participants of our user group that scored for each of the prototypes and that results highlight variations in user satisfaction across prototypes, through 6 scales which are Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation, and Novelty. Based on the UEQ scores as an initial quantitative measurement, we compare the usability of the prototypes by performing more comparative statistical analysis.

When collecting the responses to the UEQ questionnaire from the users we have included random attention check questions as shown in figure 5.5, to avoid the random guesses of careless answers, ensuring that the users have rated the questions after carefully reading them with attention. It helps to ensure the reliability of the responses, which leads to performing accurate statistical analysis.

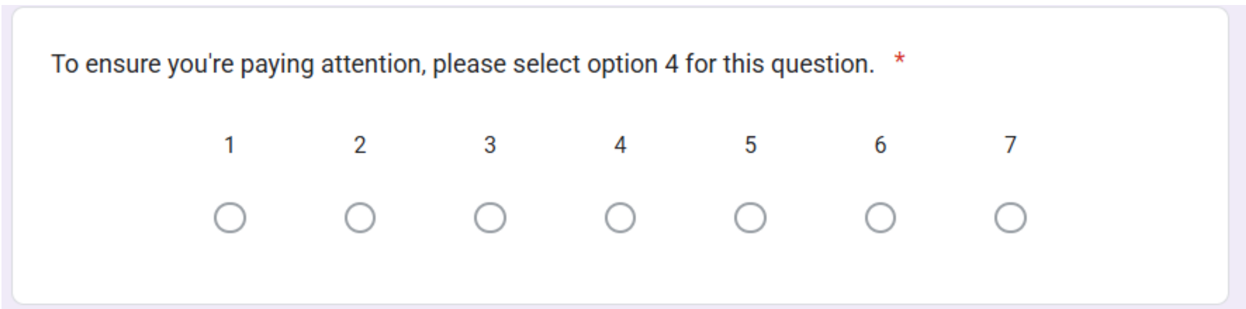


Figure 5.5: Attention check questions for UEQ

5.3 Effectiveness of the System

To compare the effectiveness of the three systems that we have developed, we were interested in two significant parameters: learning outcomes and user engagement. These criteria were chosen to assess how well each system helped students with ADHD in both maintaining attention during the learning process and achieving their academic goals.

5.3.1 Learning Outcomes - Quiz Performance Analysis

To measure learning outcomes, each participant completed a standardized quiz after using the respective system. The purpose of these tests was to evaluate students' understanding and retention of the information covered during the learning session. The quiz scores gave a quantitative indicator of each system's ability to support effective learning among students with ADHD.

All three systems had quizzes with similar questions and followed an identical structure to ensure fairness and consistency across evaluations. Each quiz consisted of 10 questions, which included 3 multiple-choice questions (MCQs), 5 fill-in-the-blanks questions and 2 true or false questions. This balanced question structure provided a comprehensive assessment of each learner's understanding of the learning content.

5.3.2 User Engagement—Task Analysis via Microsoft Clarity

User engagement was assessed through detailed task analysis using Microsoft Clarity, which is a behavioral analytics tool that captures real-time user interaction data and provides insights into how users interact with web-based platforms. Through features such as session recordings, heatmaps, click tracking, and scroll depth analysis, we were able to observe how users navigated each system.

The analysis revealed differences in engagement levels among the three systems. Metrics such as session duration, active interaction rates, and signs of user frustration such as rage clicks or repeated back-and-forth navigation were used to assess how users engage with the interfaces to evaluate the effectiveness of the learning management systems.

For our study we integrated Microsoft Clarity into each of the three prototypes to monitor the user behavior during the evaluation sessions. To ensure privacy we used secure recording through Clarity where user inputs were masked. Some key features of Clarity used in our analysis included:

- Session recordings, which allowed us to watch how users interacted with different elements of the interfaces
- Heatmaps, which provided a visual representation of frequently viewed and clicked areas on each screen
- Scroll depth analysis, which helped determine how far users navigated through the content
- Click tracking, to monitor where users clicked, how often, and in what sequence
- Rage click detection, which identified repeated, rapid clicks on the same element, often a sign of user frustration or confusion.

5.4 Results

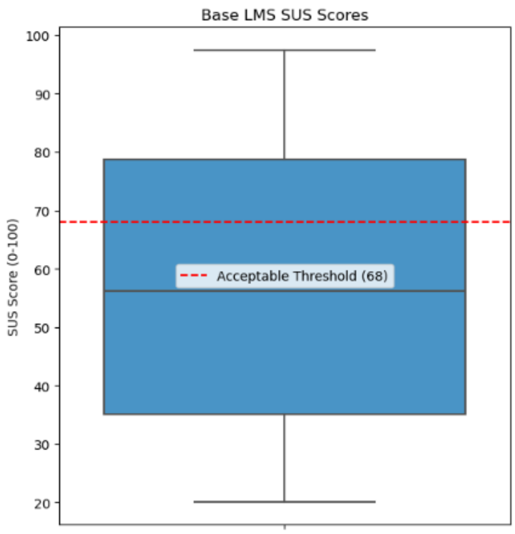
This section presents a comprehensive analysis and interpretation of the results obtained through the evaluation of three different Learning Management System (LMS) prototypes: Base LMS (Prototype 1), Manually Improved LMS (Prototype 2), and AI-Generated LMS (Prototype 3). First we have gathered data through SUS, UEQ responses, quiz scores representing the learning outcomes, user feedbacks, interviews and Microsoft clarity results for each of the LMS and then based on them performed advanced comparative analysis to derive final results. Like that, the evaluation focuses on usability, user experience, and effectiveness using standardized instruments and statistical methods to assess which LMS prototype is most suitable for individuals with ADHD.

5.4.1 Experiment 1 - Evaluation of Prototype 1 : Base LMS

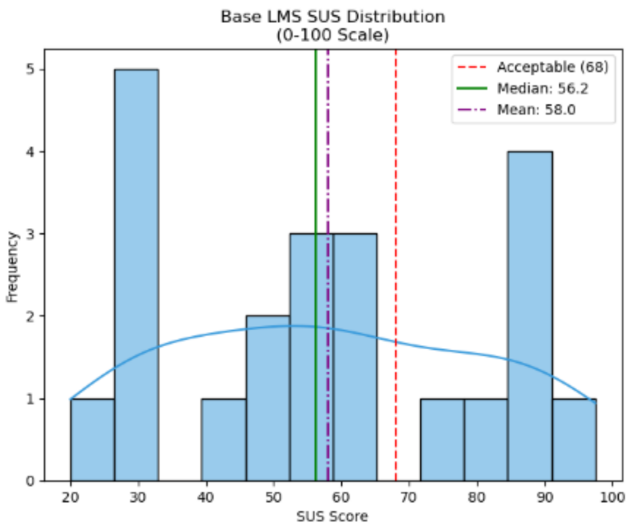
Through the first experiment we evaluated the usability and effectiveness of the base LMS prototype, designed for a general audience. Usability was assessed using the SUS and the UEQ. Further, Effectiveness of the prototypes was assessed by considering users' learning outcomes through their quiz scores and Microsoft clarity results.

SUS Scores Analysis

The SUS scores provide a standardized measure of system usability. The average System Usability Scale score is considered as 68 [55]. According to the boxplot in figure 5.6(a), Base LMS median SUS score is below the acceptable threshold(68), representing the need of improvements. It displays a wide spread in SUS scores, with many users rating the system below the median value. As well as the wide inter quartile range indicates high variability of the SUS scores distribution and overall poor usability.



(a) Base LMS SUS Scores Boxplot



(b) Base LMS SUS Scores Distribution

Figure 5.6: Base LMS SUS Score visualizations: (a) Boxplot and (b) Distribution.

Base LMS SUS Score Distribution in figure 5.6(b) gives the median is 56.2 and mean is 58, both are below the acceptable threshold of 68. It shows that most participants scored the base LMS under 60, indicating that the usability of base lms is significantly poor.

Further, we have categorized the SUS scores of Base LMS according to the interpretation of [55], scores which are equal to or higher than 80.3 (80.3-100) as Excellent, equal to or higher than 68 (68-80.3) as Acceptable usability, equal to or higher than 51 (51-68) as Marginal usability and below 51 as Poor usability. According to the figure 5.7, the results confirm that the majority of participants (40.9%) rated the system as poor usability by scoring below 51. Only 22.7% participants have rated it as a system of excellent usability.

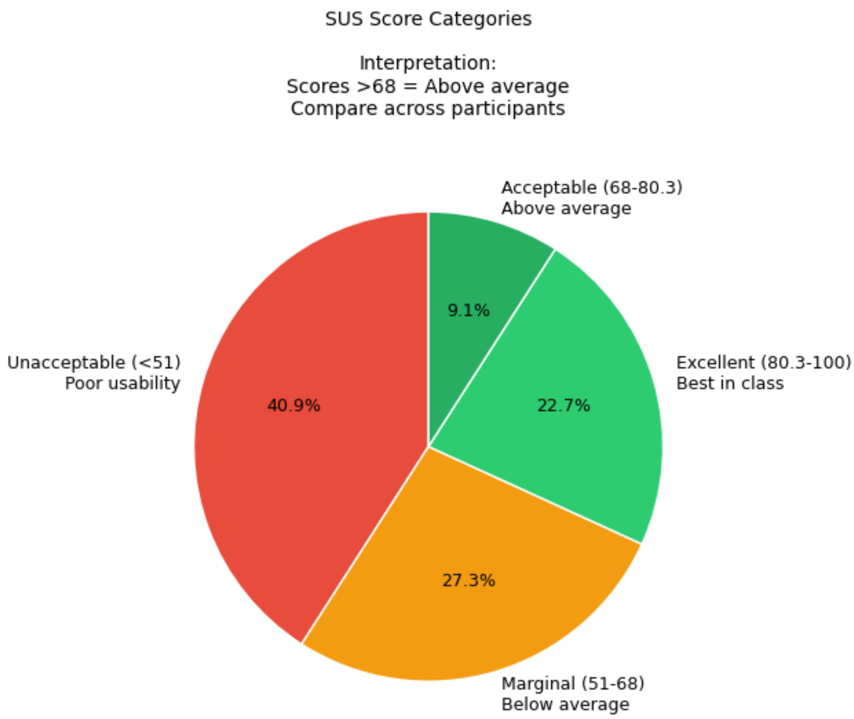


Figure 5.7: Base LMS SUS Scores Categories

UEQ Results Analysis

The seven scale UEQ questionnarie was used to assess the usability of Base LMS across six dimensions. The boxplot of Base LMS UEQ scores in figure 5.8, indicates that median UEQ score is below than the acceptable threshold of 4.0. It also shows low UEQ scores with high variance. We can identify some outliers with lower scores below Q1, which may lead to inconsistencies and more variance of the distribution.

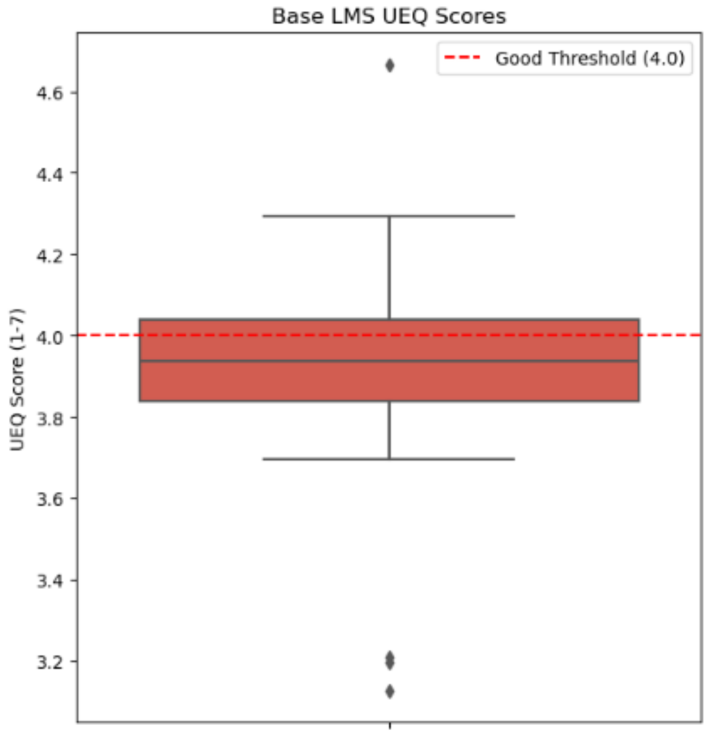


Figure 5.8: Base LMS UEQ Scores Boxplot

According to the studies [48] based on the interpretation of standard deviations of UEQ scales, three thresholds have been defined as,

- High agreement: Standard deviation of scale below 0.83.
- Medium agreement: Standard deviation of scale between 0.83 and 1.01.
- Low agreement: Standard deviation of scale above 1.01.

The figure 5.9 indicates the mean scores and standard deviatons of Base LMS across the six UEQ dimensions. It shows that lowest mean scores in the dimensions of Stimulation (3.60) & Novelty (3.58) of Base LMS. Only Dependability (4.19) & Perspicuity (4.07) have acquired mean scores higher than 4.0. The Radar chart represents an uneven profile, indicating inconsistent UX.

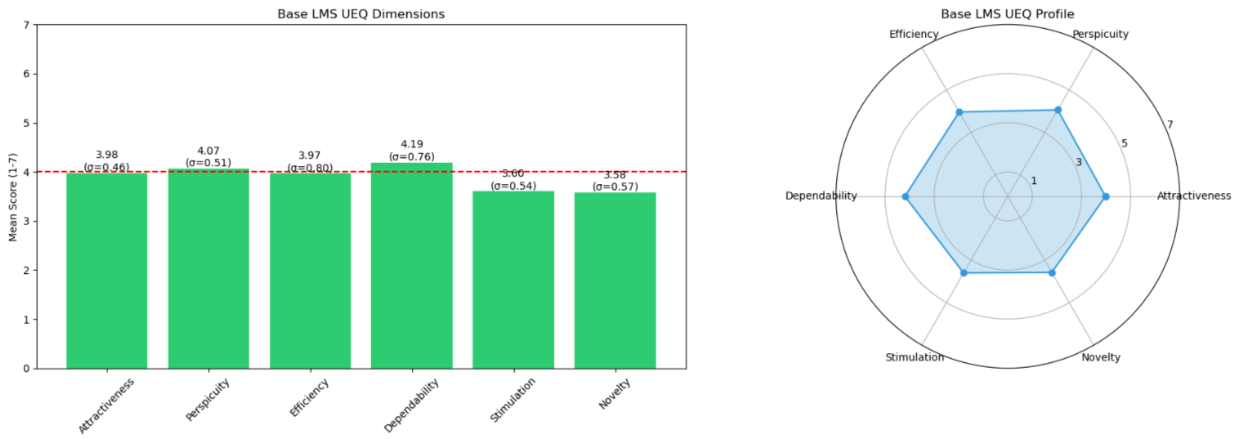


Figure 5.9: Base LMS UEQ Dimensions

Further, the figure 5.10, clearly represents high, medium, low agreement dimensions based on the standard deviation. It indicates only Dependability (4.19) & Perspicuity (4.07) have positive scores greater than 4, representing that all the other four dimensions needed to be improved in Base LMS. All the six dimensions indicates a high agreement by accquiring standard deviation below 0.83. It shows that users have a consistent perception of the UX aspect ensuring the reliability and confidence in the results.

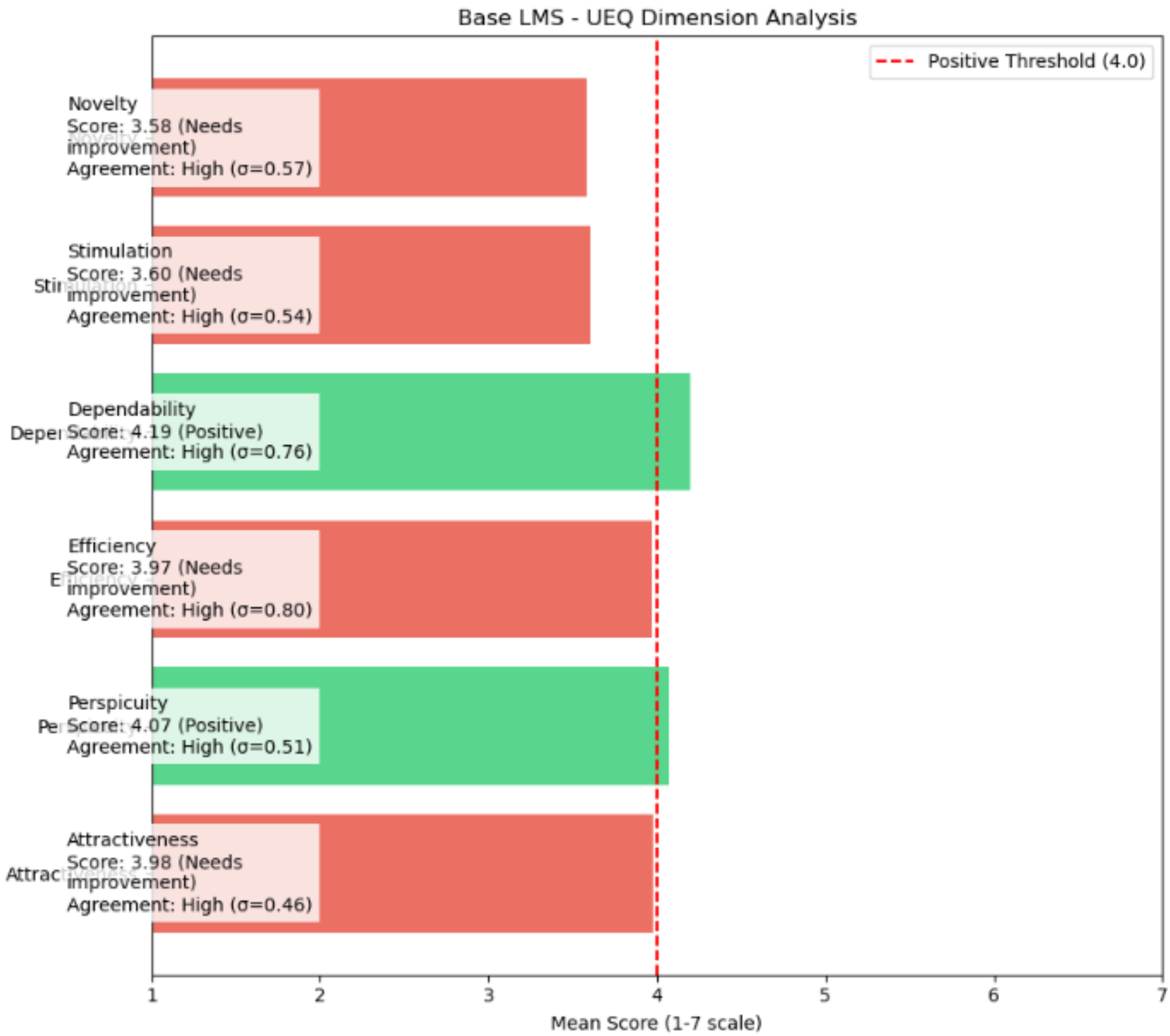


Figure 5.10: Base LMS UEQ Dimension Analysis

Learning Outcomes

To measure the learning outcomes, we considered the quiz scores acquired by the individual participants, allowing them to answer a similar set of questions in each of the LMS. The boxplot of figure 5.11, indicates low mean quiz scores and a wide interquartile range, implying high variance, indicating inconsistency in learning outcomes.

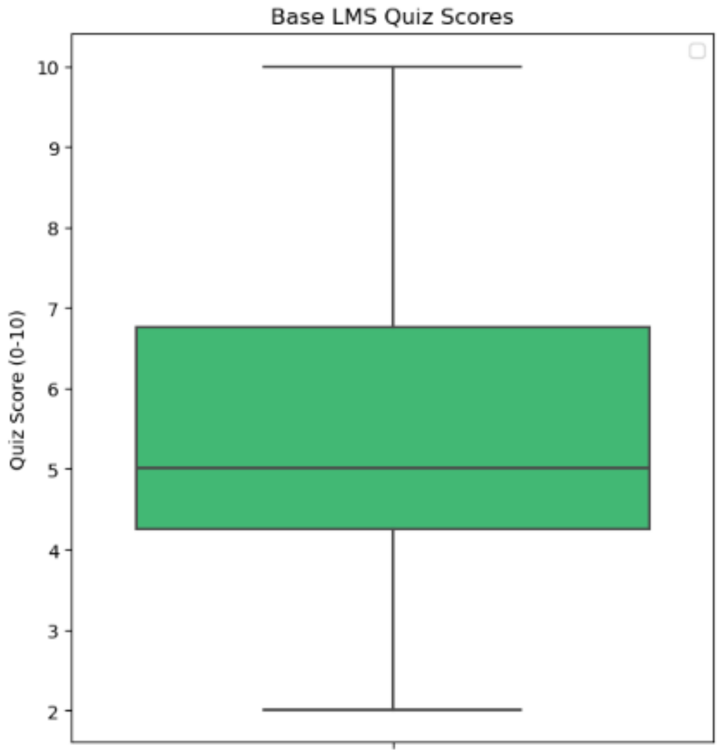


Figure 5.11: Base LMS Quiz Scores Boxplot

Further, the figure 5.12, indicates the mean as 5.4, showing that most users have acquired average marks, not much higher marks, suggesting that users struggled to retain and apply knowledge. It shows a normal unimodal distribution.

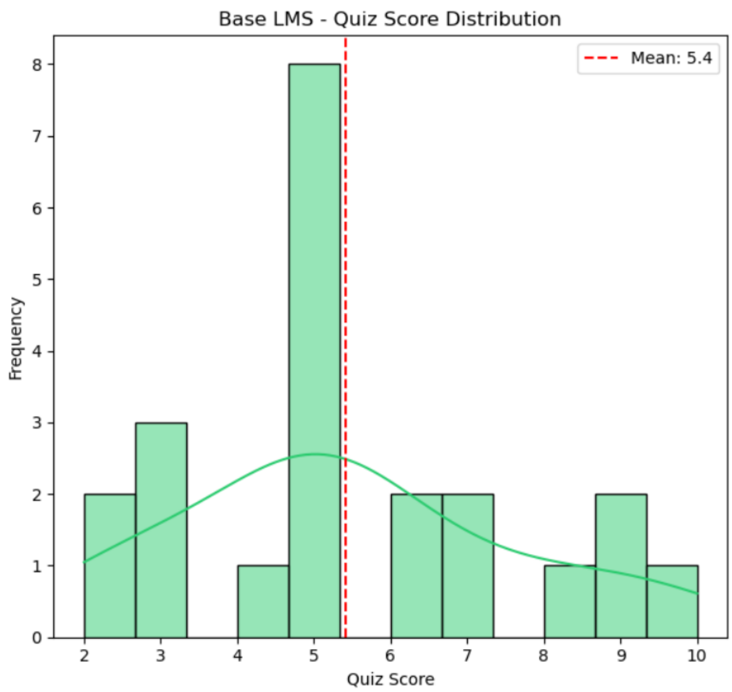


Figure 5.12: Base LMS Quiz Scores Distribution

Microsoft Clarity results

To assess the effectiveness of the Base LMS, a combination of behavioral metrics from Microsoft Clarity and visual evidence from the attention heatmap was analyzed. This evaluation reflects how users with ADHD interacted with the system and identifies areas of strength and potential improvement.

1. Attention Map Insights

Figure 5.13: Base LMS Clarity Attention Map

The attention heatmap of the pages revealed that:

- High attention (red/orange zones) was concentrated at the top and central parts of the page, especially around section headers and visually prominent text areas like "Distinguishing Science from Pseudo-Science" and "Key Principles of Scientific Theories." in the content page.
- Medium engagement (yellow/green zones) extended across the middle content, indicating that users skimmed or briefly read these sections.
- Lower engagement (blue zones) appeared toward the bottom of the pages.

This gradient suggests that while users were initially engaged, attention gradually dropped, likely due to cognitive fatigue or layout density.

2. Key Microsoft Clarity Metrics

Metric	Value	Interpretation
Pages per session	6.70 (avg)	Indicates good exploration across the LMS, suggesting that users were willing to browse through multiple pages.
Scroll depth	90.40% (avg)	Shows that users scrolled through almost the entire page, but the heatmap indicates that attention reduced toward the end.
Active time spent	2.1 min out of 7.6 min total	Only 28% of the time was spent actively engaging. This implies significant passive time – likely due to distraction or cognitive overload.
Rage clicks	2.78%	Low but noteworthy – some interface elements have been frustrating or unclear to a few users.
Dead clicks	26.85%	High rate suggests users clicked on non-interactive or poorly responsive elements, indicating usability issues.
Quick backs	36.11%	High rate implies that users frequently left pages quickly, possibly due to confusion, dissatisfaction, or lack of clarity.
Excessive Scrolling	0%	No sessions recorded excessive scrolling, which may indicate content is well-structured or concise.

Table 5.2: Key Microsoft Clarity Metrics - Prototype 1

3. Performance Score and Technical UX

Performance score: 77/100 from 118 pageviews

- 45.8% Good
- 54.2% Needs Improvement
- 0% Poor

Breakdown:

- Largest Contentful Paint - LCP (7.6s): Poor – content took too long to become visible, likely impacting user patience and initial attention.

- Interaction to Next Paint - INP (250ms): Needs improvement – interaction feedback was slightly delayed, possibly leading to dead or rage clicks.
- Cumulative Layout shift - CLS (0.005): Good – layout stability was strong, no disruptive shifting during use.

Base LMS displayed several strengths, including high scroll depth and an acceptable number of pages per session, showing that users were focused to explore the content. Additionally, the low number of rage clicks suggested minimal frustration, and the structure and content appeared engaging enough to capture early attention.

However, some significant flaws and weaknesses were identified. The users were active only for 2.1 minutes despite longer session times, pointing to potential attention deficits or passive browsing. Also, high dead click rates and frequent quick backs highlighted issues with intuitiveness and interface clarity.

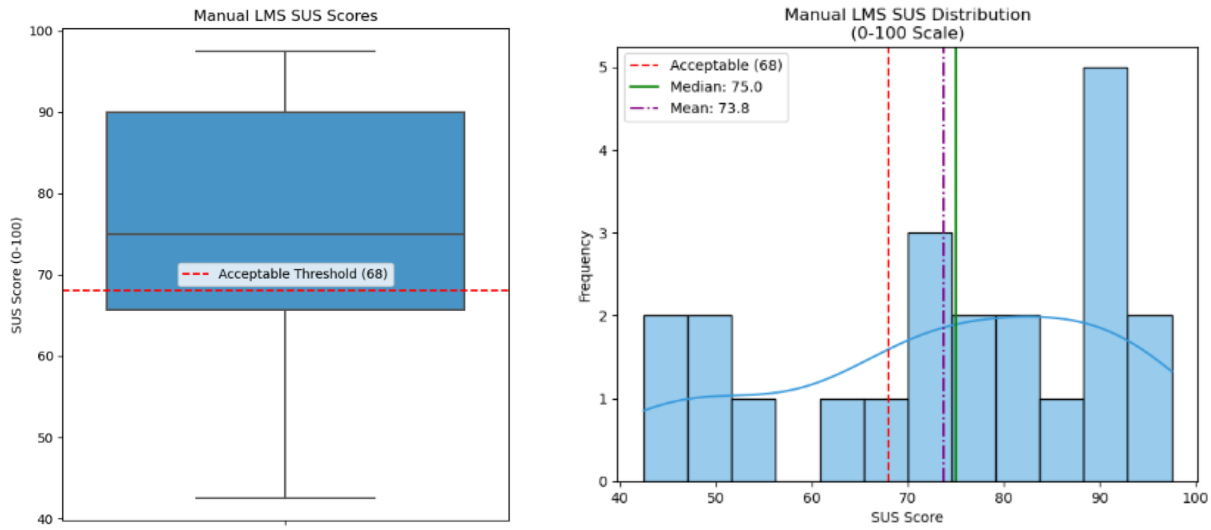
Performance related concerns, such as slow loading times (notably poor LCP scores), most likely caused interference with the user experience by affecting first impressions and the ability to maintain focus.

5.4.2 Experiment 2 - Evaluation of Prototype 2 : Manually Improved LMS

In the second experiment, ADHD participants used the manually improved LMS prototype tailored to address ADHD features. This version incorporate design adjustments based on findings from Experiment 1, such as optimized layouts, reduced cognitive load, and enhanced focus supporting features. The usability of this prototype was also evaluated using SUS, UEQ. And the effectiveness of the Manual LMS prototype was assessed by considering users' learning outcomes through their quiz scores and Microsoft Clarity results.

SUS Scores Analysis

SUS score was used to measure the system usability of Manual LMS prototype interacting with the same user group. According to the boxplot in figure 5.14(a), Manual LMS's median SUS score is above the acceptable threshold(68). The scores have shifted upwards compared to the Base LMS. The median score has improved and the spread has narrowed slightly.



(a) Manual LMS SUS Scores Boxplot (b) Manual LMS SUS Scores Distribution

Figure 5.14: Manual LMS SUS Score visualizations: (a) Boxplot and (b) Distribution.

Manual LMS SUS Score Distribution in figure 5.14(b) gives the median as 75 and the mean as 73.8, both are above the acceptable threshold of 68. It is left-skewed and shows that significant amount participants scored the Manual LMS above 68, indicating that the usability of Manual lms is comparatively higher than Base LMS.

Further when considering the categorical view of SUS by the figure 5.15, the results confirm that the majority of participants (36.4%) rated the system as excellent usability by scoring above 80.3, showing a significantly improved usability, but some users still struggled.

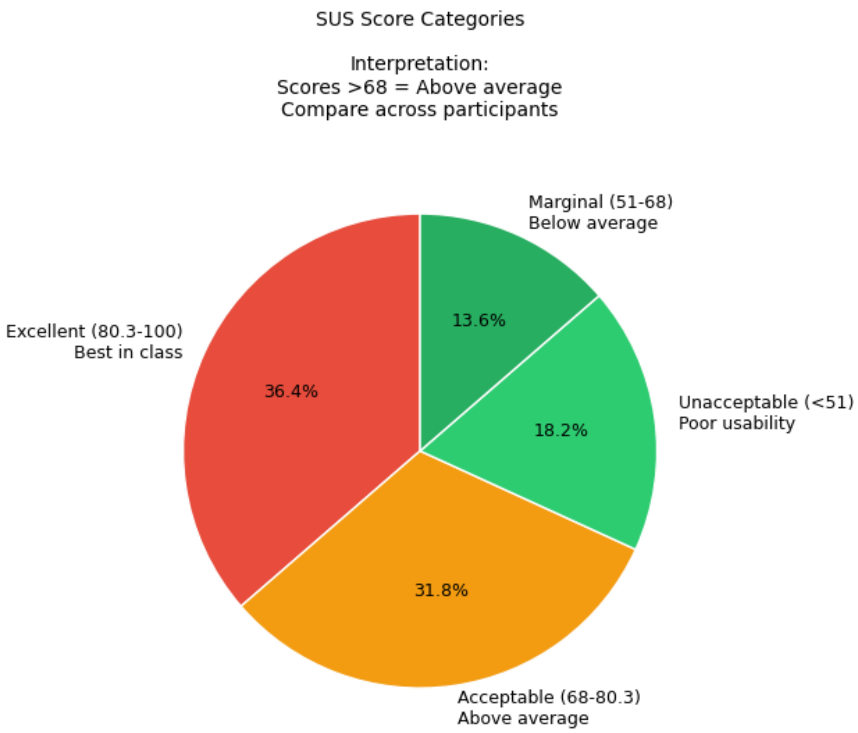


Figure 5.15: Manual LMS SUS Score Categories

UEQ Results Analysis

UEQ questionnarie was used to assess the usability of Manual LMS across six dimensions. The boxplot of Manual LMS UEQ scores in figure 5.16, indicates that median UEQ score is still below than the acceptable threshold of 4.0. It also shows low UEQ scores with high variance. No outliers can be identified and UEQ scores of Manual LMS show noticeable improvement over Base LMS.

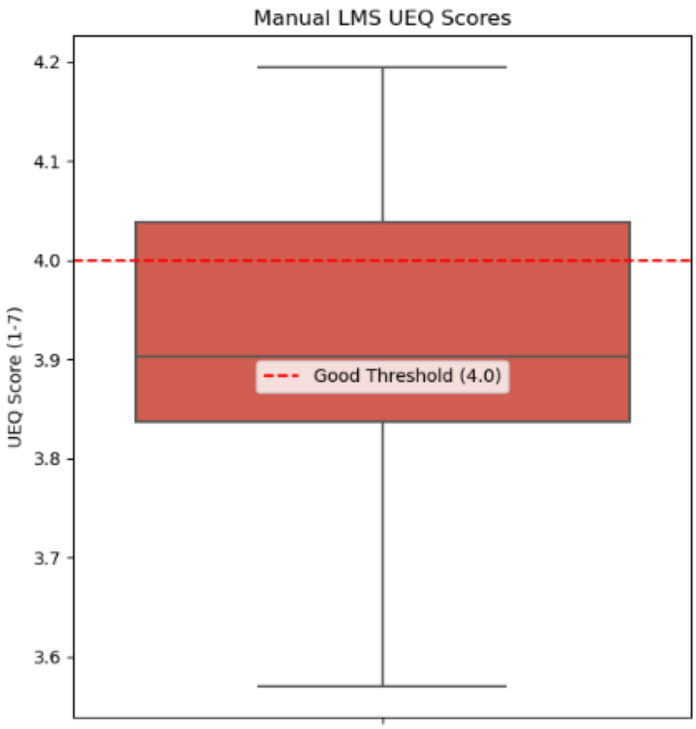


Figure 5.16: Manual LMS UEQ Scores Boxplot

The figure 5.17 indicates the mean scores and standard deviatons of Base LMS across the six UEQ dimensions. It shows that lowest mean scores in the dimension of Novelty (3.76) of Manual LMS. Still now only Dependability (4.07) & Perspicuity (4.09) have acquired mean scores higher than 4.0. Radar Chart seems to be more balanced but still lacks stimulation & Novelty.

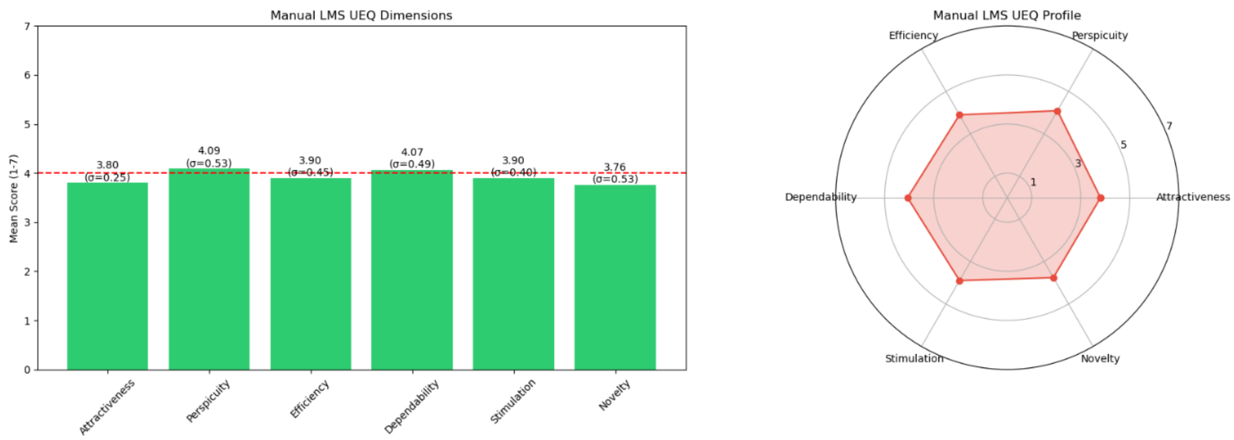


Figure 5.17: Manual LMS UEQ Dimensions

Further the figure 5.17, clearly represents high, medium, low agreement dimensions based on the standard deviation. It indicates only Dependability (4.07) & Perspicuity (4.09) have positive scores greater than 4, representing that all the other four dimensions still relatively low. All the six dimensions indicates a high agreement by accquiring standard deviation below 0.83. It shows that users have a consistent perception of the UX aspect ensuring the reliability and confidence in the results.

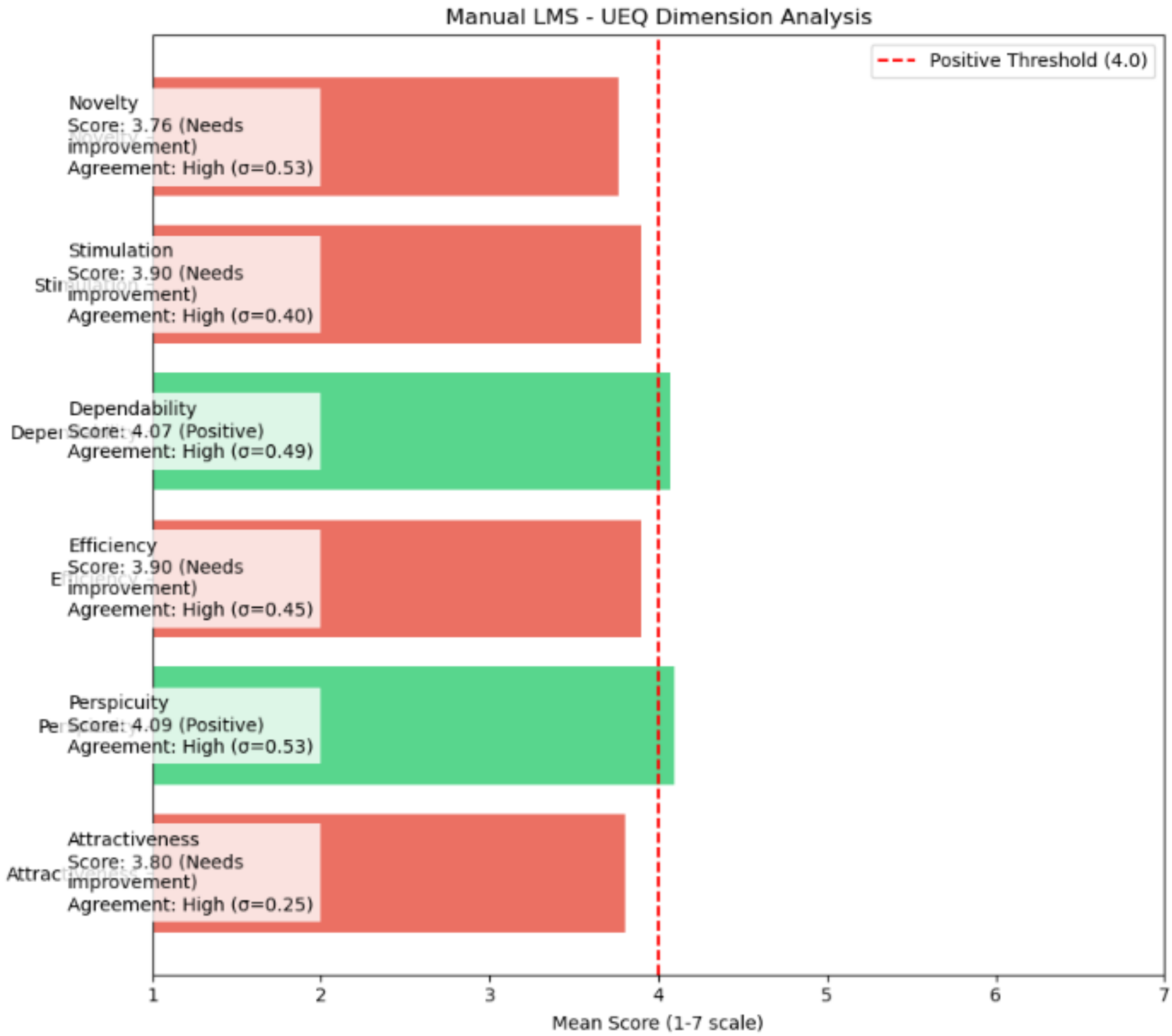


Figure 5.18: Manual LMS UEQ Dimension Analysis

Learning Outcomes

For measuring the learning outcomes of Manual LMS we considered the quiz scores acquired by the individual participants. The boxplot of figure 5.19, indicates an improved median quiz score and reduced variance compared to Base LMS.

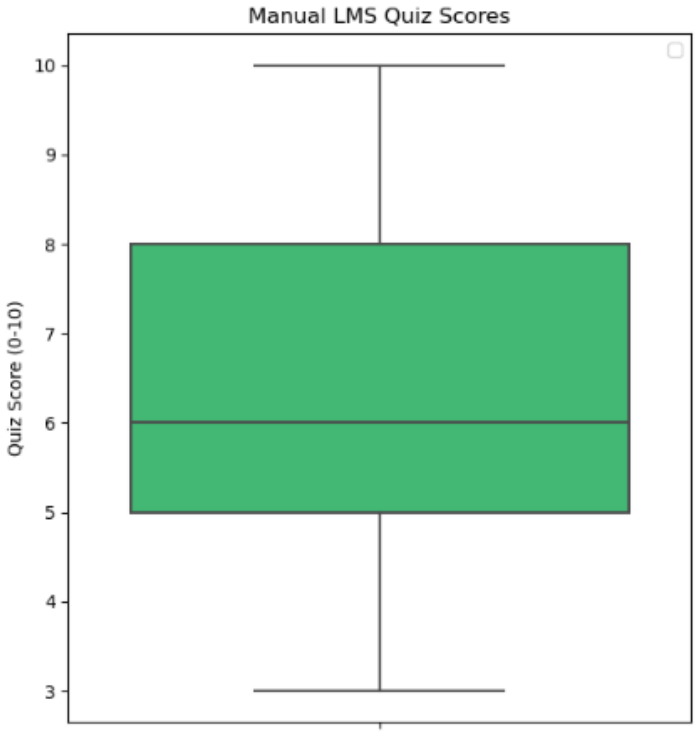


Figure 5.19: Manual LMS Quiz Scores Boxplot

Further, the figure 5.20, indicates the mean is 6.5, which is comparatively higher than the value acquired in Base LMS. The distribution is roughly normal and unimodal. When compared to Base LMS quiz score Manual LMS shows improved comprehension and retention.

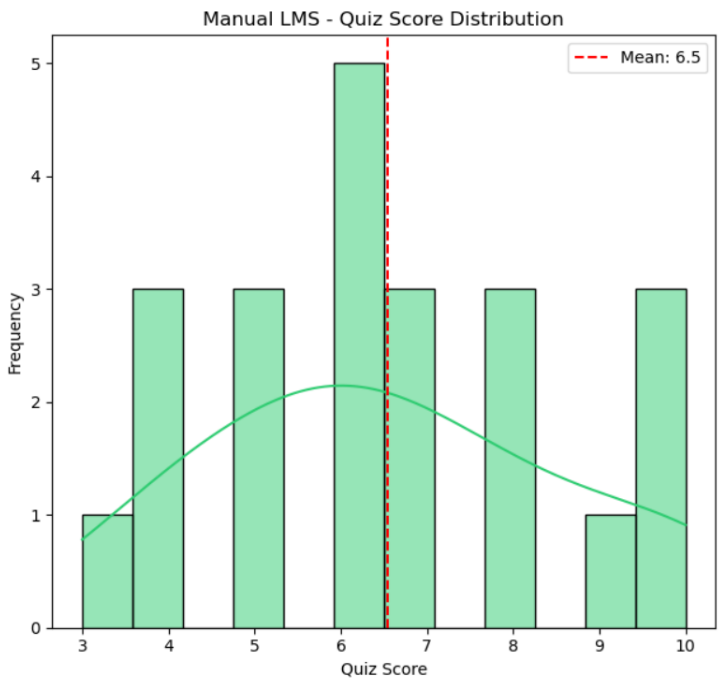


Figure 5.20: Manual LMS Quiz Scores Distribution

Microsoft Clarity Results

1. Attention Map Insights

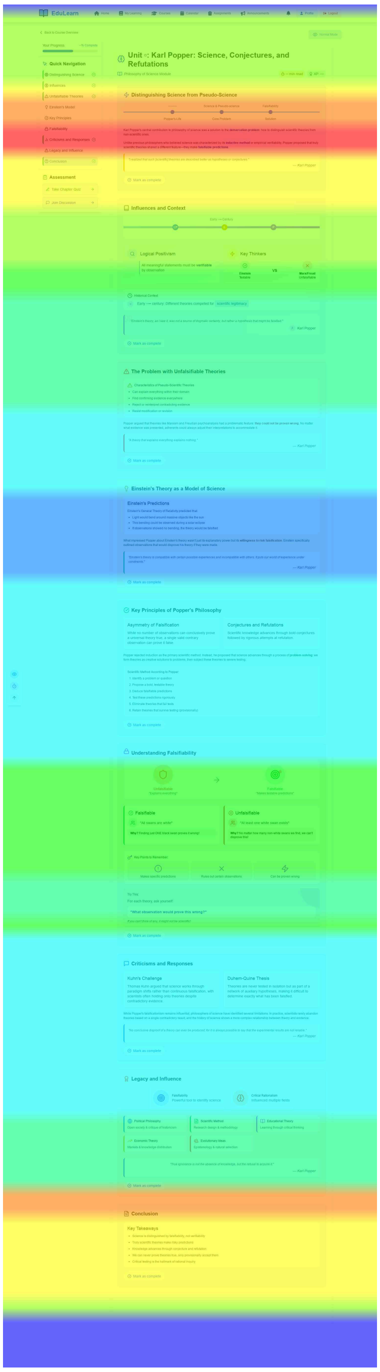


Figure 5.21: Manual LMS Clarity Attention Map

The heatmap revealed a more balanced engagement across the page compared to the base LMS:

- High attention (red/orange zones) appears not only at the top but also around some sections in the middle of the pages.
- Mid-level engagement (yellow/green) extends evenly throughout the page.
- Unlike the Base LMS, this layout retained attention in the middle and lower sections, indicating successful re-engagement techniques for ADHD users.

This suggests that strategic structuring, use of interactive or visual content, and section-based layouts helped sustain focus across the entire page. It indicated better pacing and cognitive engagement compared to Prototype 1, with more users navigating deeper into the page, although attention still dips slightly in content-dense areas.

2. Key Microsoft Clarity Metrics

Metric	Value	Interpretation
Pages per session	8.62 (avg)	Shows users were navigating between sections, not stuck or dropping off. Users browsed more pages, suggesting higher engagement.
Scroll depth	85.34% (avg)	Users almost fully explored each page indicating it as very ADHD-friendly. Layout encouraged full reading.
Active time spent	4.5 min out of 11.2 min total	40% active engagement time. Might indicate that visual cues grabbed attention but didn't always translate into active engagement.
Rage clicks	0%	Suggests users didn't feel frustrated or confused by any interactive elements indicating highly usable and ADHD-considerate interface design.
Dead clicks	11.76%	Much improved over Prototype 1 (26.85%) – Indicates better User Experience and smoother click guidance.
Quick backs	23.53%	Lower than base LMS – suggests users were less likely to abandon a page prematurely indicating better content clarity and navigational flow. Positive retention indicator. Shows reduced cognitive overload or bounce-back frustration.
Excessive Scrolling	0%	No sessions recorded excessive scrolling, which may indicate content is well-structured or concise.

Table 5.3: Key Microsoft Clarity Metrics – Prototype 2

3. Performance Score and Technical UX

Performance score: 82/100 from 71 pageviews

- 51.3% Good
- 48.7% Needs Improvement
- 0% Poor

Breakdown:

- Largest Contentful Paint - LCP (6.9s): Still poor, though slightly improved – large visual elements may slow initial load.
- Interaction to Next Paint - INP (190ms): Better – quicker interaction feedback, supporting smoother engagement.
- Cumulative Layout shift - CLS (0.003): Excellent layout stability – helps avoid attention breaks caused by shifting content. Important for users who struggle with shifting UI elements.

Prototype 2 showed improvements in user engagement and interface usability. Active engagement time rose to 4.5 minutes out of an 11.2-minute average session, indicating better attentional retention and more meaningful interaction. Learners demonstrated high scroll depth (85.34%) and strong content exploration, averaging 8.62 pages per session. The complete absence of rage clicks points to a frustration-free experience, while reduced dead clicks and fewer quick backs suggest enhanced interface clarity and smoother navigation.

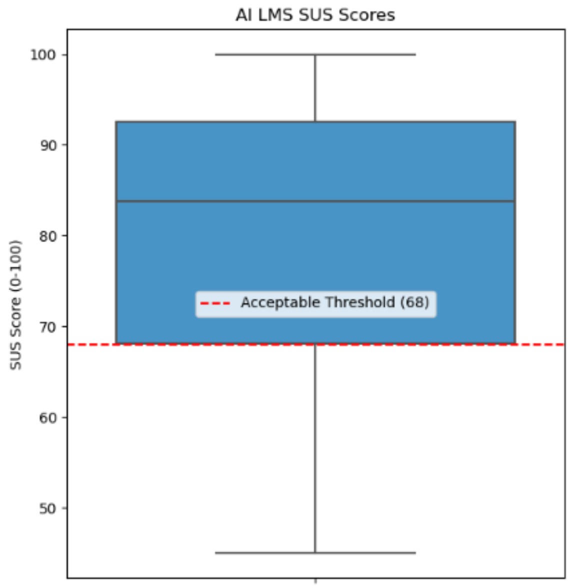
Heatmap analysis revealed consistent attention distribution across content sections, indicating that design pacing, layout, and micro-interactions effectively supported sustained focus. However, active engagement still made up less than half of the total session time, highlighting an opportunity to boost focus further. Additionally, while the 11.76% dead click rate reflects progress, it still hints at potential confusion around certain interface elements. Performance issues, such as slow loading times (LCP at 6.9s) and moderate interaction delays (INP at 190ms), may also disrupt flow during attention shifts which is a critical concern for ADHD-sensitive users.

5.4.3 Experiment 3 - Evaluation of Prototype 3 : AI-Generated LMS

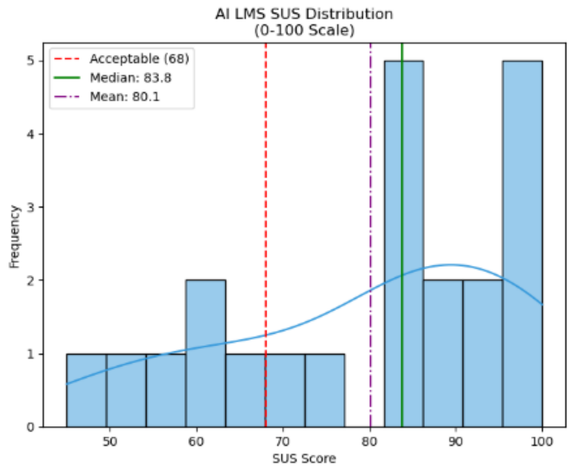
The third experiment focus on evaluating the AI-generated LMS prototype. ADHD participants evaluated this version, which integrated generative AI-driven designs informed by user feedback and HCI heuristics. Usability evaluations were again involved SUS, UEQ analysis. Further Effectiveness of the AI LMS prototype was assessed by considering users' learning outcomes through their quiz scores and Microsoft clarity results.

SUS Scores Analysis

The SUS score was used to measure the system usability of AI LMS prototype interacting with the same user group. According to the boxplot in figure 5.22(a), AI LMS median SUS score is above the acceptable threshold(68). Shows the highest median SUS score among the three LMS platforms. Scores are tightly clustered at the top end.



(a) AI LMS SUS Scores Boxplot



(b) AI LMS SUS Scores Distribution

Figure 5.22: AI LMS SUS Score visualizations: (a) Boxplot and (b) Distribution.

AI LMS SUS Score Distribution in figure 5.22(b) gives the median is 83.8 and mean is 80.1, both are above the acceptable threshold of 68. Most ratings fall within the 80–100 range. It is left-skewed and shows that significant amount participants scored the AI LMS above 68, indicating that the usability of AI LMS is significantly improved than Base LMS and Manual LMS.

Further, when considering the categorical view of SUS by the figure 5.23, the results confirm that the most of participants (63.6%) rated the AI LMS as excellent usability by scoring above 80.3 indicating a strong usability preference. It shows a significantly improved usability of AI LMS compared to Base LMS and Manual LMS.

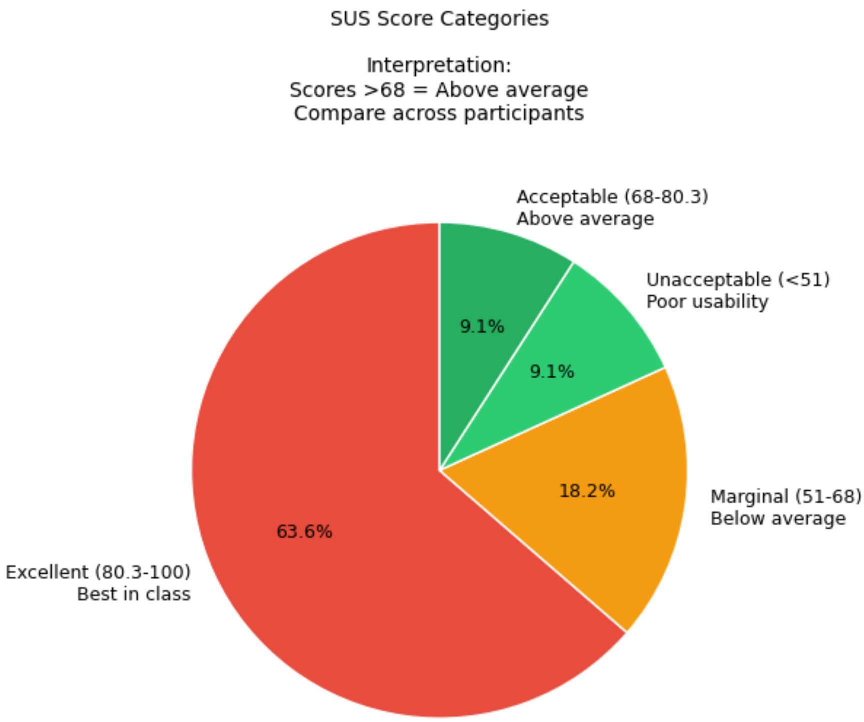


Figure 5.23: AI LMS SUS Score Categories

UEQ Results Analysis

UEQ questionnarie was used to assess the usability of AI LMS across six dimensions. The boxplot of AI LMS UEQ scores in figure 5.24, indicates that median UEQ score is still below than the acceptable threshold of 4.0. UEQ scores are high and tightly distributed.

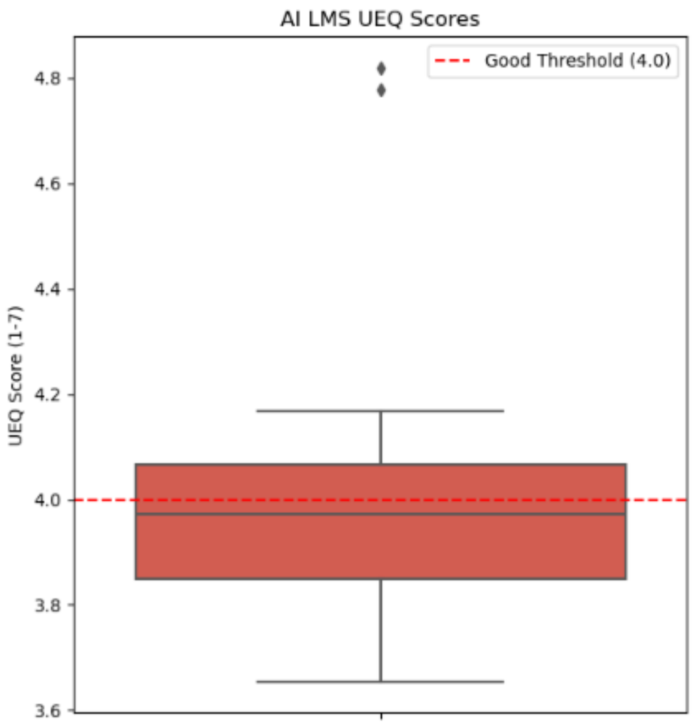


Figure 5.24: AI LMS UEQ Scores Boxplot

The figure 5.25 indicates the mean scores and standard deviatons of AI LMS across the six UEQ dimensions. It shows that lowest mean score in the dimension of Attractiveness (3.80) of AI LMS. Dependability (4.31) & Novelty (4.07) have acquired mean scores higher than 4.0. Scores of the AI LMS are comparatively higher in all dimensions—especially in Novelty. Radar Chart seems to be more balanced but some lack in Attractiveness.

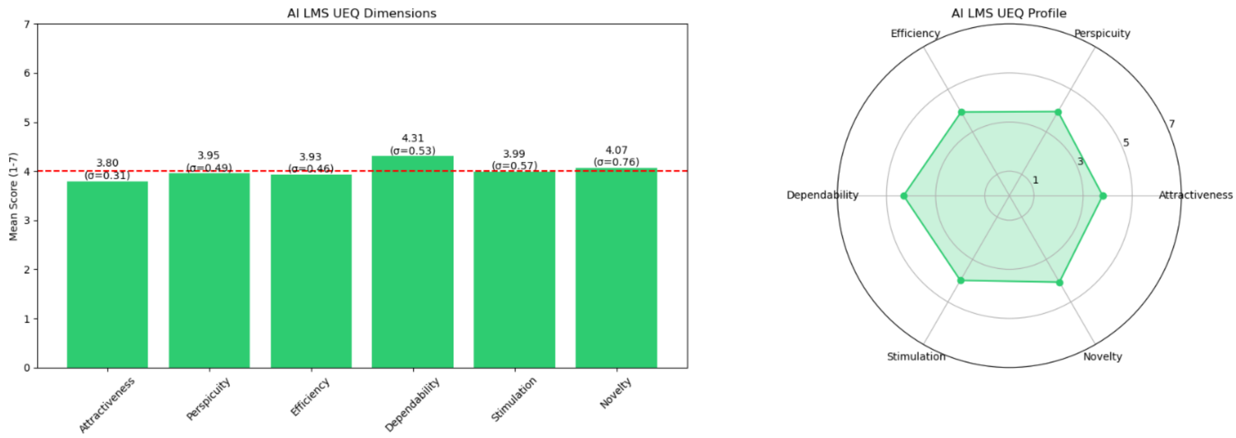


Figure 5.25: AI LMS UEQ Dimensions

Further, the figure 5.26, clearly represents high, medium, low agreement dimensions based on the standard deviation. It indicates only Dependability (4.31) & Novelty (4.07) have positive scores greater than 4. All the six dimensions indicates a high agreement by accquiring standard deviation below 0.83. It shows that users have a consistent perception of the UX aspect ensuring the reliability and confidence in the results. AI LMS outperforms both other prototypes significantly in terms of usability.

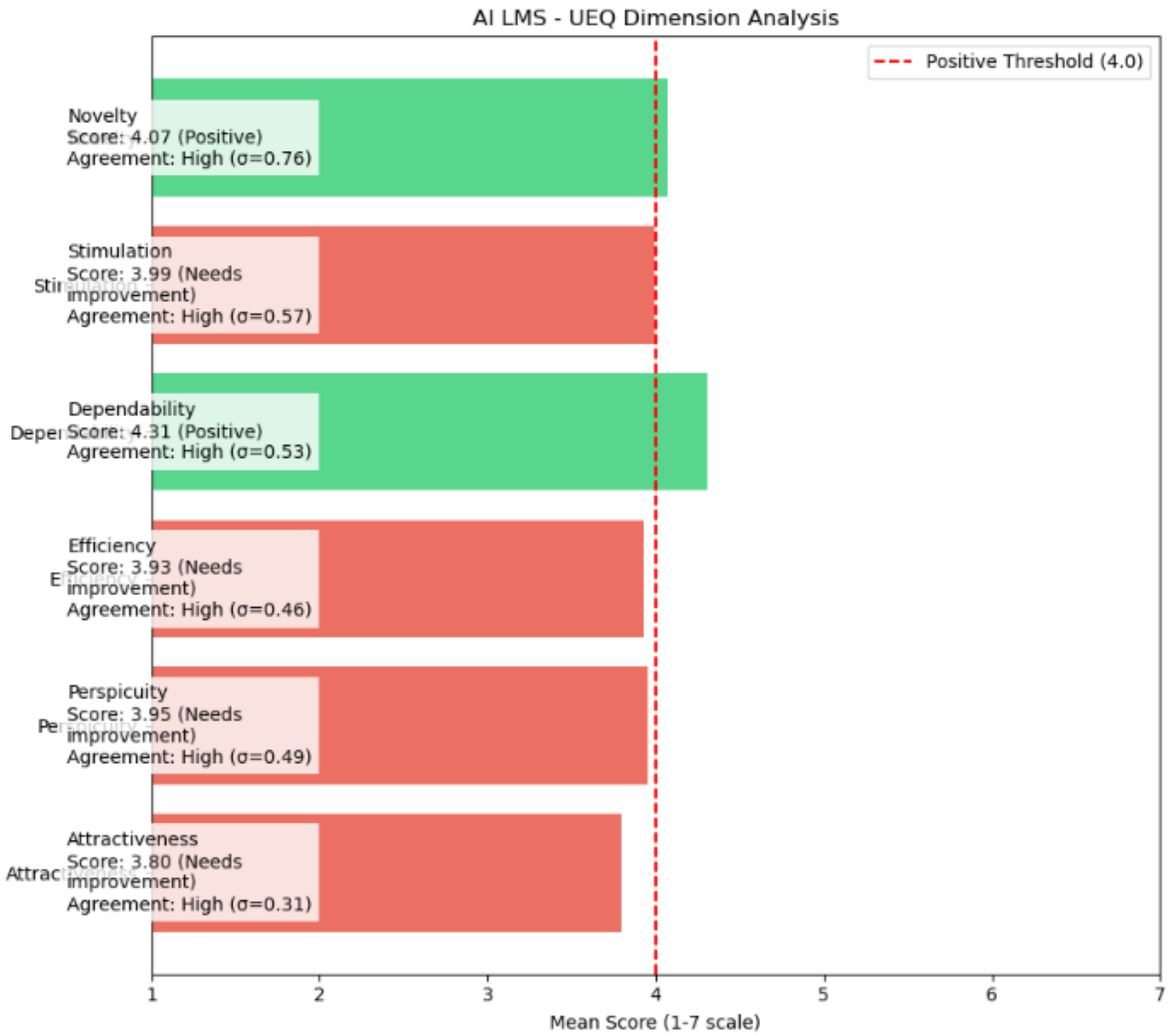


Figure 5.26: AI LMS UEQ Dimension Analysis

Learning Outcomes

To measure the learning outcomes of AI LMS we considered the quiz scores acquired by the individual participants. The boxplot of figure 5.27, indicates an improved median quiz score of median 6 higher. It represents least variance compared to Base LMS and AI LMS implying that users performed consistently well.

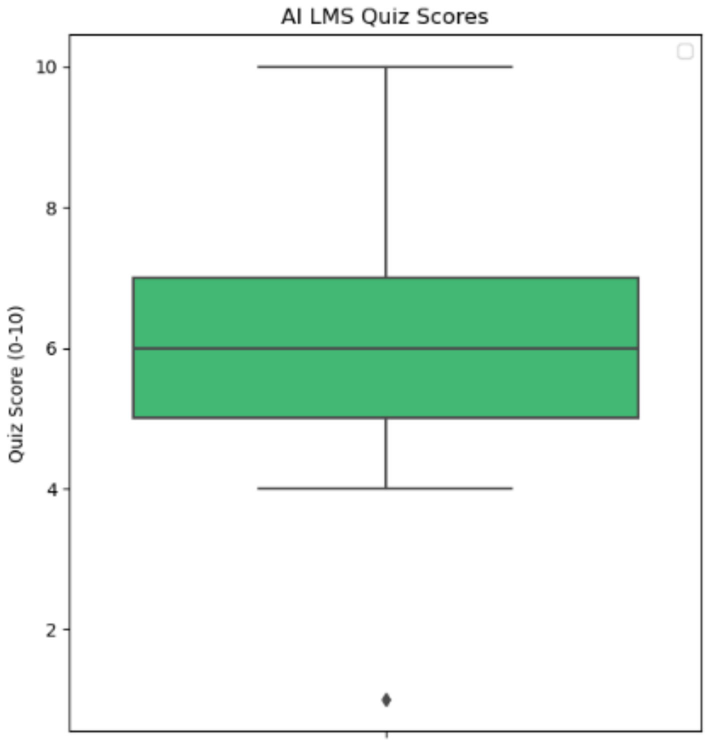


Figure 5.27: AI LMS Quiz Scores Boxplot

Further, the figure 5.28, indicates the mean is 5.9 which is comparatively higher than the value acquired in Base LMS and slightly lower than Manual LMS. The distribution is roughly normal and unimodal. When compared to Base LMS quis score AI LMS shows improved comprehension and retention.

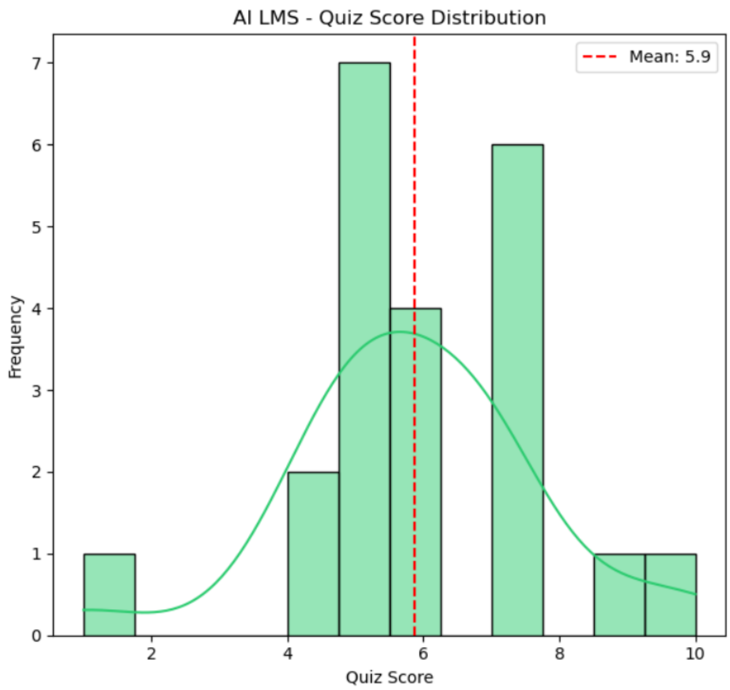


Figure 5.28: AI LMS Quiz Scores Distribution

Microsoft Clarity Results

1. Attention Map Insights

By analysing the quantitative data about user engagement gained through Microsoft Clarity, we were able to gain important insights about ADHD user behaviour with learning management systems. These behavioural patterns inferred from metrics indicate mixed engagement.

- High scroll depth metrics indicate that users were visually exposed to nearly all content, suggesting successful layout structure and pacing.
- Despite the adaptive capabilities, quick backs and high dead click rates suggest potential confusion or friction, pointing to issues with interface intuitiveness or content pace.

2. Key Metrics

Metric	Value	Interpretation
Pages per session	8.52 (avg)	High browsing activity remains consistent—users were actively engaged in exploring.
Scroll depth	94.54% (avg)	High scroll coverage indicates content layout and AI-enhanced flow are effective.
Active time spent	1.6 min out of 7.4 min total	21.6% active time shows solid, though slightly reduced, engagement from Prototype 2.
Rage clicks	4%	A slight increase may reflect frustration with unpredictable AI interactions.
Dead clicks	32%	A high score indicates that the affordances are ambiguous or that there might be an overload of AI-generated elements.
Quick backs	56%	A high percentage of users frequently exited pages rapidly, which may indicate cognitive fatigue or inadequate content relevance.
Excessive Scrolling	0%	No sessions recorded excessive scrolling, which may indicate content is well-structured or concise.

Table 5.4: Key Microsoft Clarity Metrics - Prototype 3

3. Performance Score and Technical UX

Performance score: 80/100 from 63 pageviews,

- 61.9% Good
- 34.9% Needs Improvement
- 3.2% Poor

Breakdown:

- Largest Contentful Paint-LCP (6.5s): Poor—Slow initial rendering can disrupt attention and cause restlessness, especially for ADHD learners.
- Interaction to Next Paint (INP) (150 ms): Good—Interaction response is fast and likely contributes positively to engagement.
- Cumulative Layout Shift—CLS (0.001): Good—Layout stability remains a strength, ensuring a visually consistent experience.

It was evident that learners were profoundly and broadly engaged with the content, as Prototype 3 demonstrated strong user exploration, excellent scroll depth, and solid pages per session. The active engagement time remained consistent with previous versions, indicating that users were attracted to the dynamic, potentially more personalized content. An ADHD-friendly experience was further enhanced by a stable visual layout (CLS) and optimized interaction response time (INP). The AI-generated adaptive content flow appeared to increase initial interest and personalization, which is noteworthy. It is important to note that the prototype also introduced new obstacles. A potential friction in the AI-driven interface, potentially resulting from cognitive overload or unpredictability, is indicated by an increase in rage clicks and a high quick back rate. The dead click rate was the highest among all prototypes, which was likely the result of dynamic elements that appeared interactive but were not, resulting in confusion. Furthermore, users with attention regulation challenges continue to encounter persistent difficulties with sluggish load times (LCP), which impede immediate engagement.

5.5 Overall System Evaluation

This section presents a comprehensive study of the comparative performance between the three LMS prototypes which are Base LMS, Manually Improved LMS, and AI-Generated LMS through advanced statistical methods. The evaluation employs multiple analytical approaches to ensure accurate, scientifically valid conclusions about system usability and effectiveness of the prototypes for finding the optimal learning interfaces for enhancing the learning experience of ADHD students.

5.5.1 Descriptive statistics Analysis

Descriptive statistics provide a foundational understanding of central tendencies (mean) and dispersion (standard deviation) of the data. Below figure 5.29 represents the descriptive statistics of the SUS scores for each LMS.

SUS Scores Descriptive Statistics:

	Base LMS_SUS score	Manual LMS_SUS score	AI LMS_SUS score
count	22.000000	22.000000	22.000000
mean	57.954545	73.750000	80.113636
std	23.345699	17.589059	17.156098
min	20.000000	42.500000	45.000000
25%	35.000000	65.625000	68.125000
50%	56.250000	75.000000	83.750000
75%	78.750000	90.000000	92.500000
max	97.500000	97.500000	100.000000

Figure 5.29: SUS Scores Descriptive Statistics

According to this figure 5.29, when considering central tendency of SUS scores, Base LMS (Mean=57.95) falls below the acceptable threshold (68) while Manual LMS (Mean=73.75) and AI LMS (Mean=80.11) are higher than the acceptable threshold. Manual LMS shows a significant improvement compared to Base LMS, while AI LMS (Mean=80.11) achieves the highest mean compared to the others. Taking into account standard deviation of SUS scores, Base LMS shows highest SD (23.35), indicating inconsistent user experiences while AI LMS has the lowest SD (17.16), demonstrating reliable performance. By analysing the range of SUS scores, Base LMS exhibits extreme scores (20-97.5) while AI LMS maintains tighter range (45-100).

According to these statistics of SUS scores, we can analyse that Base LMS with the lowest mean, indicates poor usability, Manual LMS shows a moderate increase in usability and AI LMS with the highest mean SUS, indicates exceptional usability and user satisfaction. Therefore the progression in mean scores with decreasing variability strongly suggests that ADHD-specific optimizations yield both higher and more consistent usability.

UEQ Scores Descriptive Statistics:

	Base_LMS_UEQ	Manual_LMS_UEQ	AI_LMS_UEQ
count	22.000000	22.000000	22.000000
mean	3.897803	3.919848	4.007652
std	0.359239	0.186970	0.290006
min	3.125000	3.570000	3.653333
25%	3.836667	3.836667	3.850417
50%	3.937500	3.903333	3.972500
75%	4.038333	4.038333	4.065833
max	4.666667	4.195000	4.820000

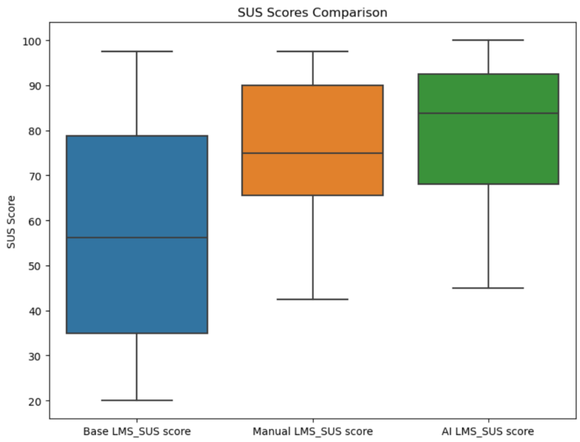
Figure 5.30: UEQ Scores Descriptive Statistics

Above figure 5.30 represents the descriptive statistics of the UEQ scores for each LMS. All systems show mean UEQ scores near neutral (4.0) but, Base LMS (3.89) and Manual LMS(3.91) fails to meet positive UX threshold. AI LMS (4.00) is the first to cross neutral benchmark. When considering the inter quartile ranges, Base LMS has widest UX perception variability while AI LMS shows most consistent positive ratings with less variance compared to other two prototypes. Therefore this results help to determine that the AI LMS scored consistently high across user experience dimensions. Further the Manual LMS improved upon the Base LMS but did not reach AI LMS levels.

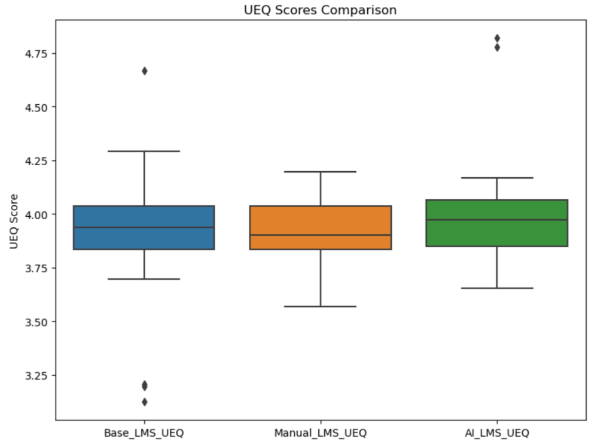
5.5.2 Comparison Visualizations Analysis

These comparative visualizations across each LMS offer an intuitive, comparative understanding of the overall system performance, supporting to finding the optimal prototype for ADHD users.

The figure 5.31(a), indicates a comparison of SUS scores across three LMS prototypes. According to this representation we can identify that the AI LMS achieved a median SUS score above 85, the Manual LMS around 72, and the Base LMS below 60. As we previously mentioned, 68 is considered as the acceptable threshold of SUS scores. Therefore this comparison indicates that AI LMS has achieved “Excellent” SUS score while Base LMS falls below 60 in the “Poor” category. According to SUS scores comparisons, AI LMS seems to achieving the highest usability.



(a) SUS Scores Comparison



(b) UEQ Scores Comparison

Figure 5.31: Comparison of SUS and UEQ Scores across prototypes.

Figure 5.31(b), indicates a comparison of UEQ scores across three LMS prototypes. Here also similar trends can be observed, with AI LMS excelling in all dimensions. However, when considering UEQ scores comparison it is not much variant across three LMS as in SUS score distribution. All the prototypes indicates median scores near the threshold while AI LMS achieves the highest score ensuring more usability.

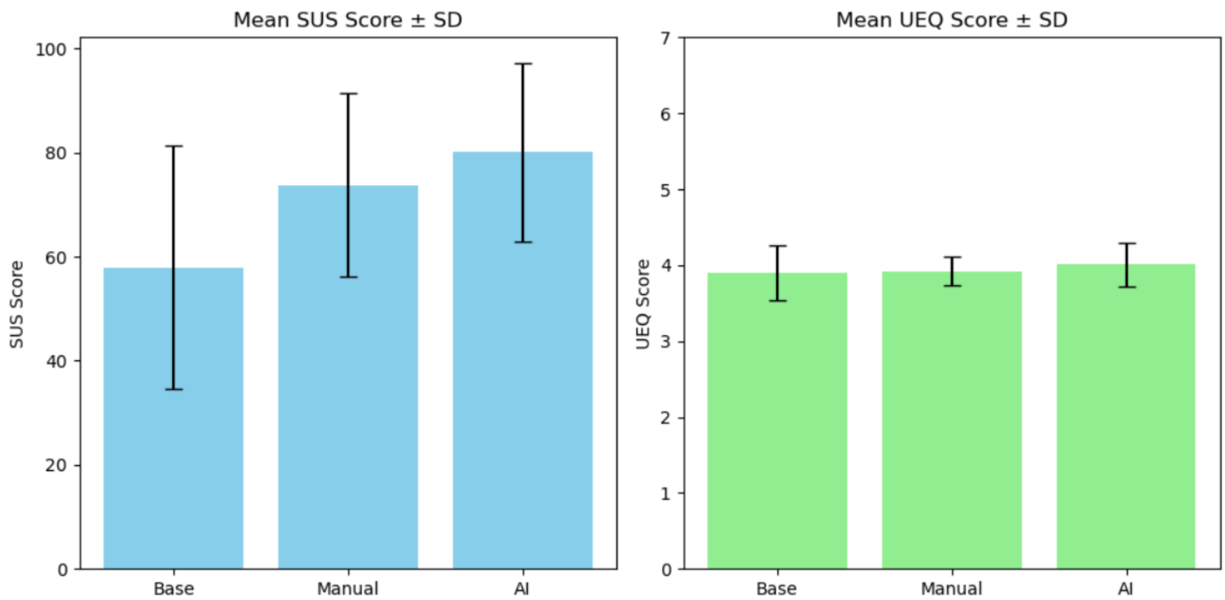


Figure 5.32: Mean SUS and UEQ Scores with Standard Deviation

Figure 5.32, represents that, the AI LMS has the highest mean and smallest standard deviation in SUS scores, implying consistent user ratings. A lower standard deviation strengthens the reliability of the system's performance across varied users. When considering the UEQ score mean values of three prototypes do not indicate a significant variability, but AI LMS mean score is slightly higher than the other two. Manual LMS standard deviation seems to be the smallest one of UEQ scores.

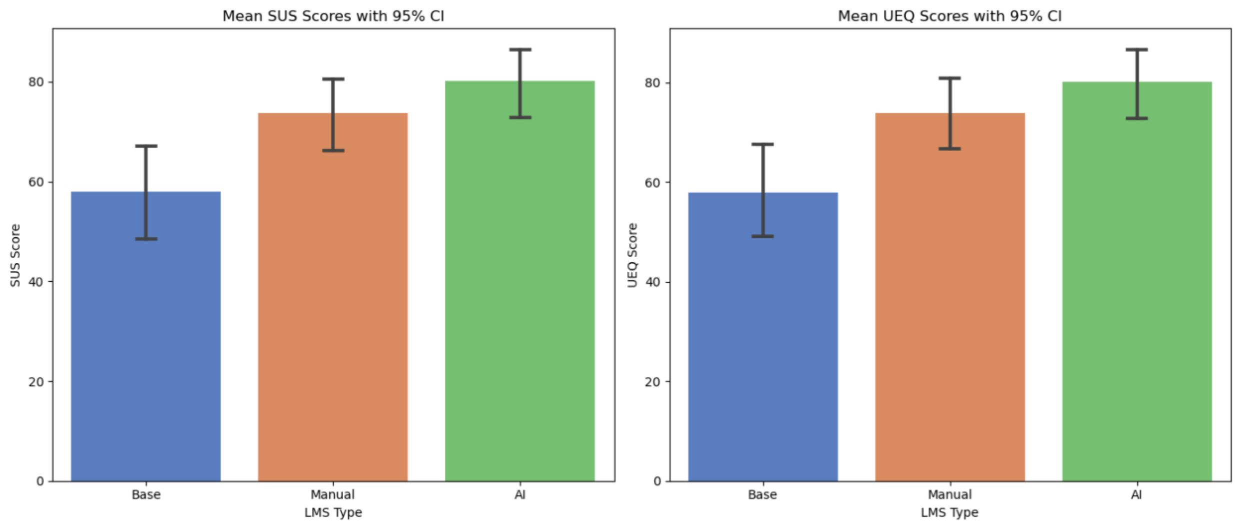


Figure 5.33: Mean SUS and UEQ Scores with 95% CI

The figure 5.33 presents the mean System Usability Scale (SUS) and User Experience Questionnaire (UEQ) scores for each LMS prototype, along with 95% confidence intervals for each mean. Confidence intervals indicate the range within which we can be 95% certain the true population mean lies, based on the sample data.

When considering the Mean SUS Scores here, Base LMS shows low (below 60) with Wide CI. It indicates low usability with high variation in user ratings, suggesting inconsistent user experiences. Manual LMS represents a moderate SUS score (around 72) with slightly narrower CI than Base LMS. It indicates improved usability over Base LMS with a more stable response trend, but still some variability exists. AI LMS shows high SUS score (above 85) with very narrow CI.It indicates that high reliability and minimal deviation among users' ratings. When considering the Mean UEQ Scores here also, Base LMS shows low score with Wide CI. Manual LMS represents a moderate SUS score with slightly narrower CI than Base LMS. And AI LMS shows high SUS score with very narrow CI. The lack of overlap between AI LMS and other prototypes' confidence intervals implies a statistically meaningful difference in usability perceptions.

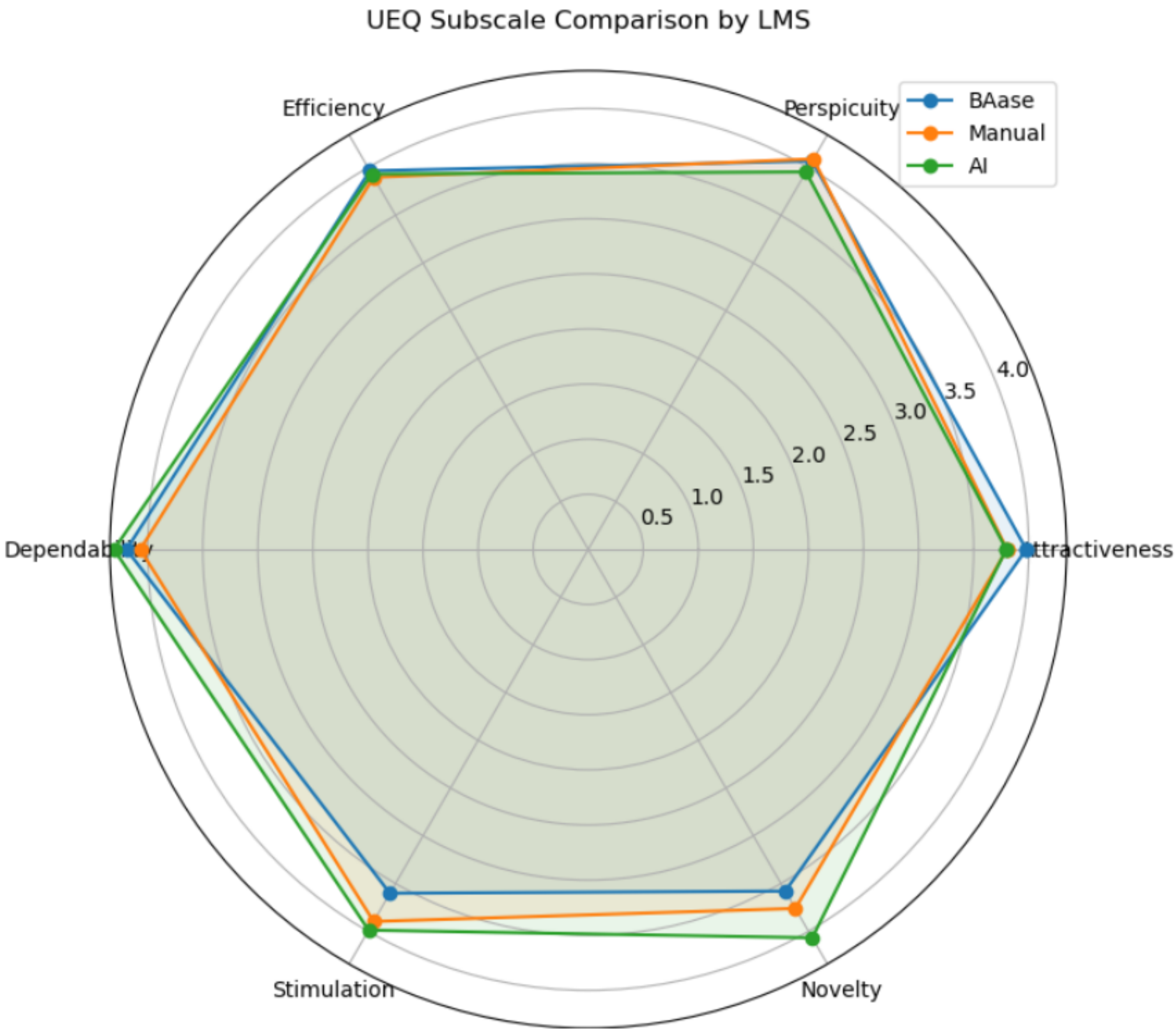


Figure 5.34: UEQ Subscale Comparison

The figure 5.34 breaks down the UEQ into its six dimensions. The AI LMS outperforms both Base and Manual LMS in almost all the dimensions including Attractiveness , Efficiency, Perspicuity , Dependability, Stimulation specially in Novelity.

5.5.3 Normality Test

Before applying inferential tests, it is necessary to check whether the data follow a normal distribution. Therefore, We performed Shapiro-Wilk Test (Normality Testing) [29] to check if the SUS & UEQ values of Base LMS, Manual LMS, AI LMS are normal or non-normal.

Normality Test Summary:

	Measure	LMS	W	p-value	Normality
0	SUS	Base	0.938631	0.185452	Normal
1	UEQ	Base	0.903014	0.034168	Non-normal
2	SUS	Manual	0.916413	0.064091	Normal
3	UEQ	Manual	0.953035	0.362282	Normal
4	SUS	AI	0.914943	0.059774	Normal
5	UEQ	AI	0.778409	0.000234	Non-normal

Figure 5.35: Normality Tests(Shapiro-Wilk)

Figure 5.35 shows the Shapiro-Wilk Test Results we derived. P-values for most of the data sets are less than 0.05, indicating that the data significantly deviate from a normal distribution. But UEQ of Base LMS and UEQ of AI LMS derived as non-normal.

We have derived below visualizations based on the normality test results to represent an in-depth analysis of System Usability Scale (SUS) score distributions using histograms and Q-Q plots. These visualizations are helpful for verifying whether the data meet the assumption of normality, which informs the choice of appropriate statistical tests parametric or non-parametric.

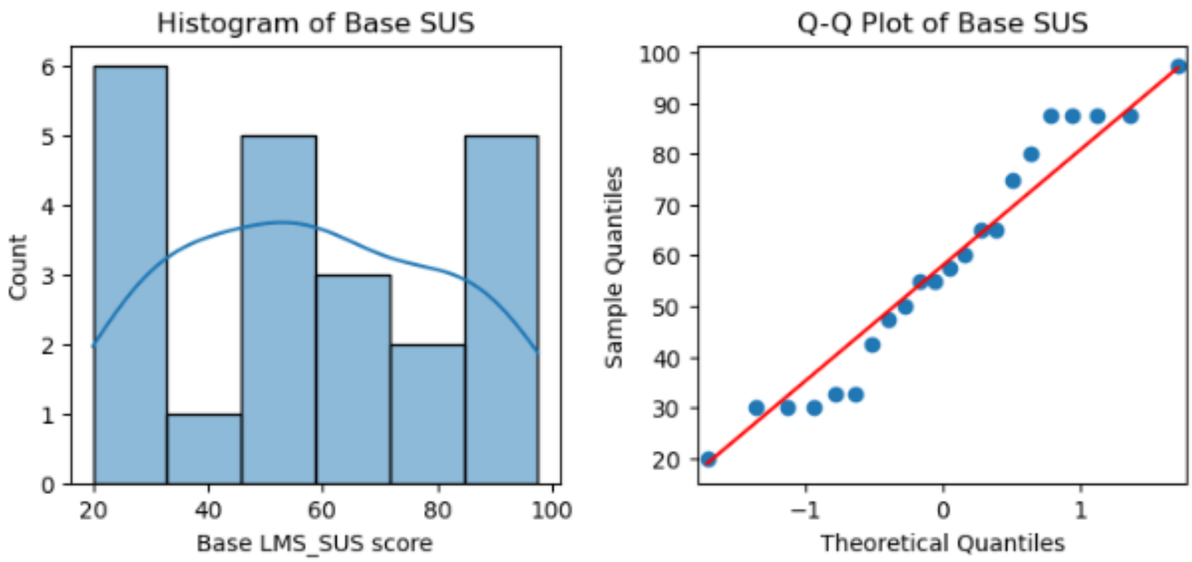


Figure 5.36: Normality Test - Histogram and Q-Q Plot for Base LMS SUS

According to the figure 5.36, histogram of Base LMS SUS appears to be approximately normally distributed. It indicates that the participants' usability perceptions of the Base LMS vary around the average, with both high and low scores but no strong skew. Q-Q Plot indicates that the points largely follow the diagonal reference line, with only minor deviations at the tails. This visually supports the assumption of normality for the Base LMS SUS scores.

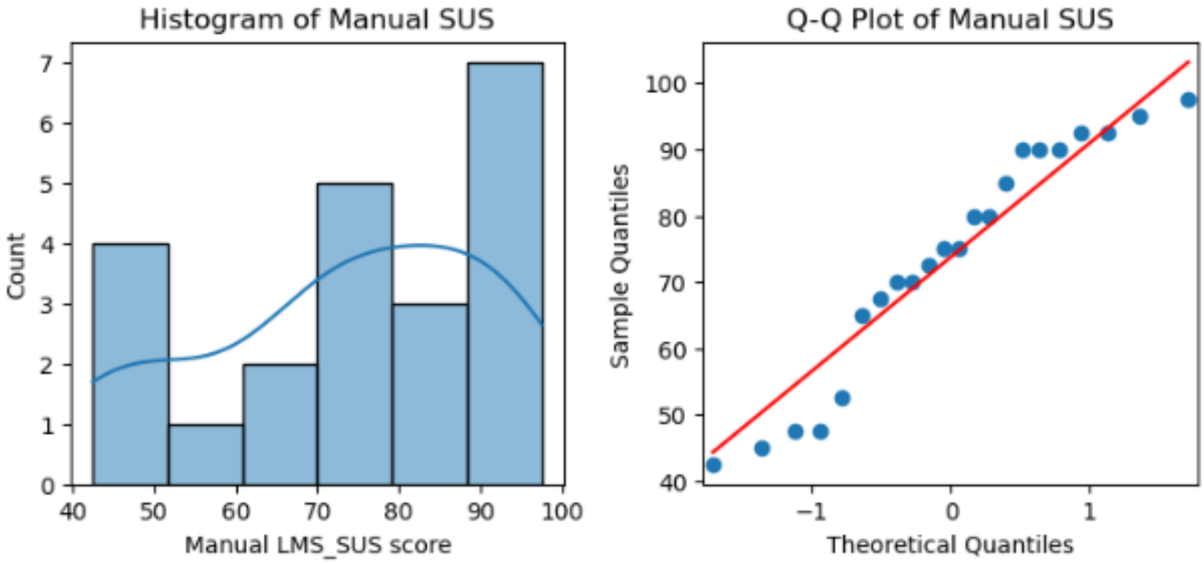


Figure 5.37: Normality Test - Histogram and Q-Q Plot for Manual LMS SUS

The figure 5.37, represents the histogram of Manual LMS SUS which is appeared to be slightly left skewed but nearly normally distributed. It indicates a concentration of high scores (above 70), with considerably fewer participants rating usability at the low end. Q-Q Plot indicates that the points deviate below the diagonal at the lower end. The presence of slightly left skew shows the departure from normality.

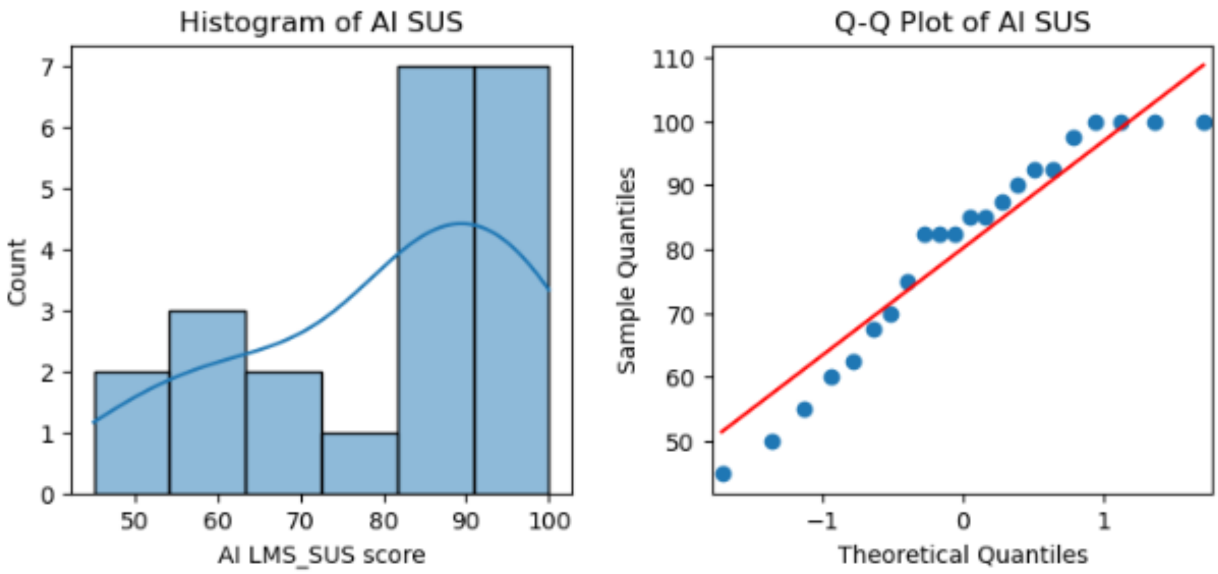


Figure 5.38: Normality Test - Histogram and Q-Q Plot for AI LMS SUS

According to the figure 5.38, histogram of AI LMS SUS appears to be clearly left-skewed (negatively skewed). Majority of scores are clustered between 85–100, with a few lower scores pulling the tail. It indicates that the AI LMS is highly rated by almost all users. However, this strong skew means the data do not conform to normality. Q-Q Plot shows that most points fall below the line at the low end. This suggests the data are not ideal for parametric testing, and non-parametric alternatives should be used.

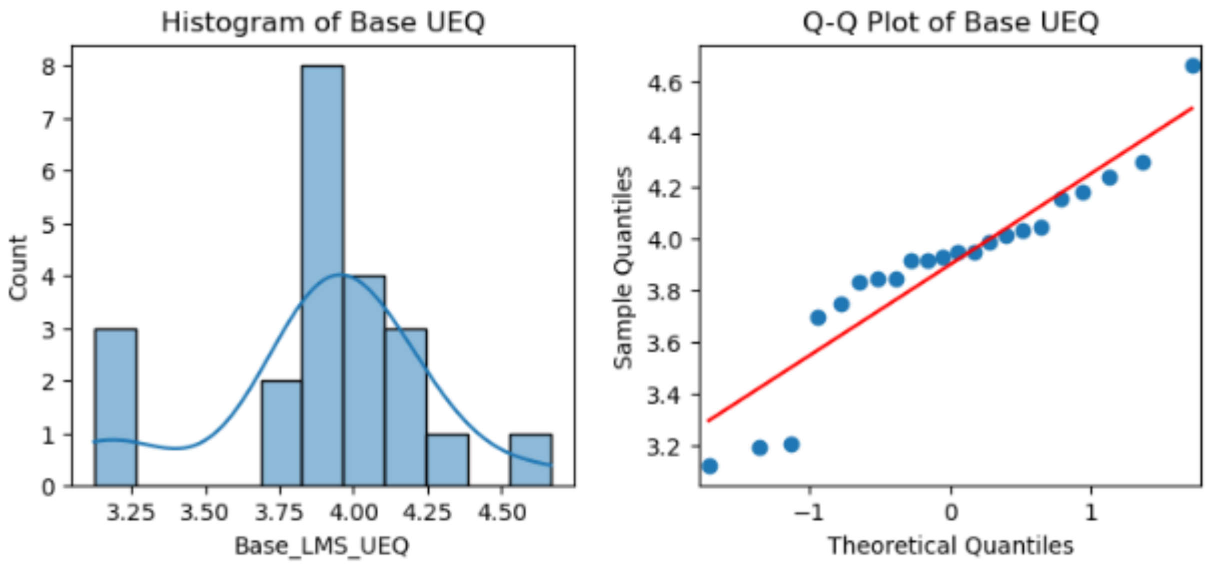


Figure 5.39: Normality Test - Histogram and Q-Q Plot for Base LMS UEQ

The figure 5.39 shows the histogram of Base LMS UEQ which is appeared to be a normal distribution. Most UEQ scores are clustered around the mean, with a relatively symmetric spread on both sides. Q-Q Plot shows that data points fall relatively close to the diagonal line, especially in the middle of the distribution. It confirms that UEQ scores for the Base LMS approximate normality. This is the most balanced of the three LMSs in terms of score distribution, though it still received the overall lowest mean scores.

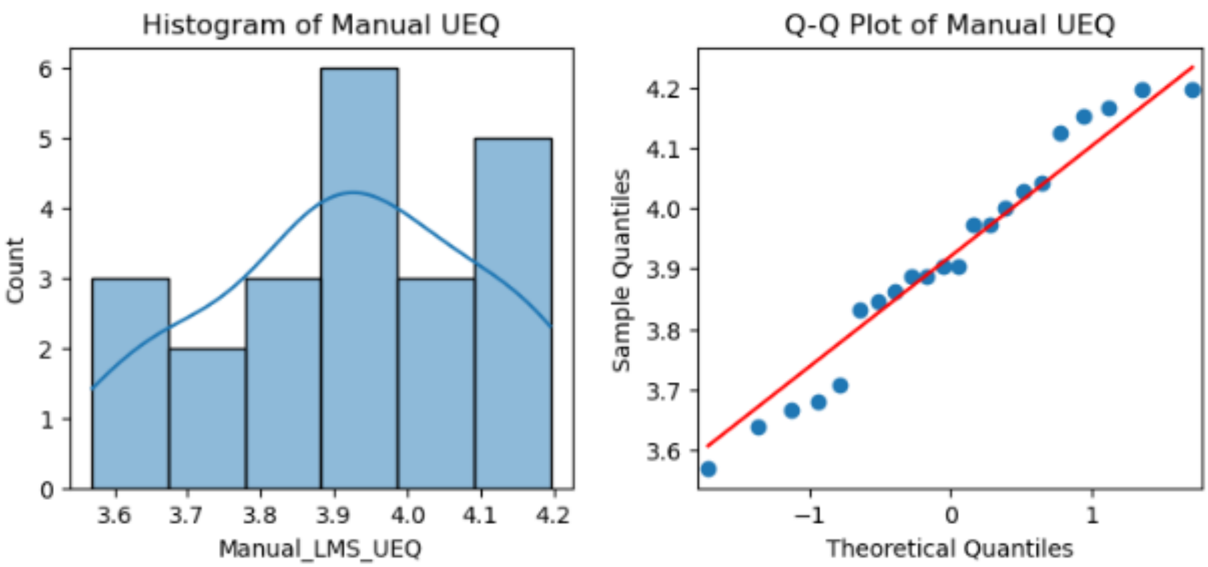


Figure 5.40: Normality Test - Histogram and Q-Q Plot for Manual LMS UEQ

According to the figure 5.40 shows the histogram of Manual LMS UEQ which indicates moderately left skewed. Most scores are concentrated in the upper-mid range (around 4–5), with fewer participants giving lower ratings. Q-Q Plot shows slight deviation from the diagonal line at both ends. Based on these it implies that use of non-parametric analysis is more appropriate for this prototype.

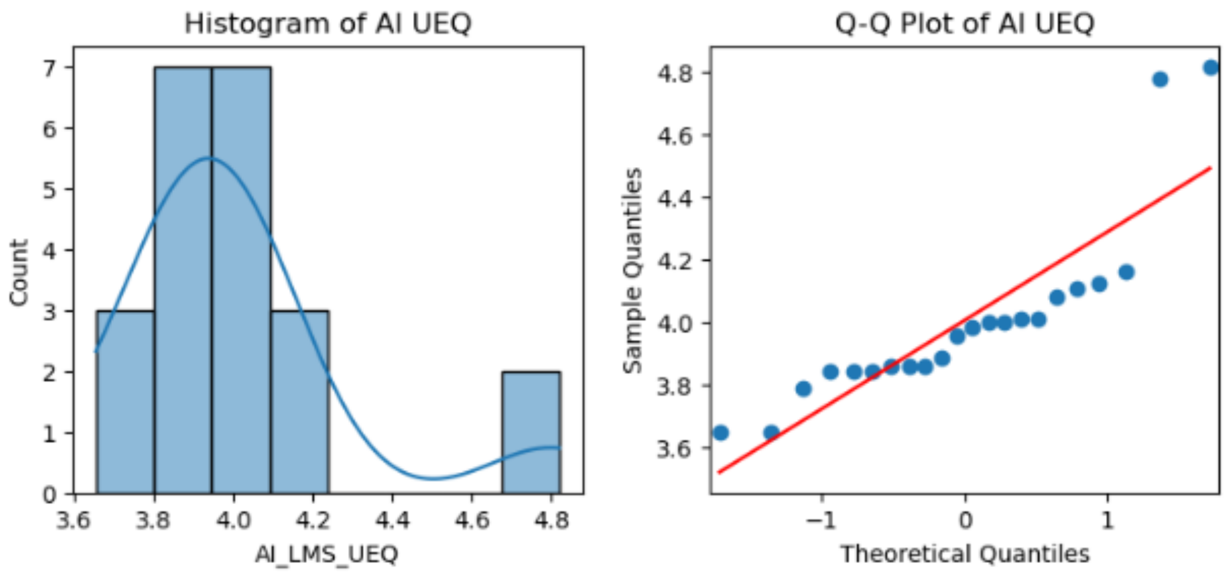


Figure 5.41: Normality Test - Histogram and Q-Q Plot for AI LMS UEQ

According to the figure 5.41 shows the histogram of AI LMS UEQ which clearly shows a right-skewed (positively skewed) distribution. Q-Q Plot also indicates the deviation from normality.

According to above Normality Test although most of the SUS & UEQ results of the LMS are normal, there include some non-normal values as well. Further in the evaluations of this study we are using repeated values by using same user group for rating all the three LMSs. One way ANOVA cannot be used for testing repeated values and non-normal values. Therefore we cannot use One way ANOVA,and as an alternative for non-parametric testing, we used Friedman Test for further analysis.

5.5.4 Friedman Test

The Friedman Test is a non-parametric statistical test [13] which can be used for analysing non-parametric, repeated-measures as an alternative to the one way ANOVA which is used for parametric testings. It evaluates whether there are statistically significant differences in user ratings across the three LMS prototypes and whether at least one system differs significantly from others.

By using Friedman test, we evaluated the statistical hypotheses:

- Null Hypothesis: There are no significant differences in SUS or UEQ scores across the three LMS variants. All systems are equally effective for ADHD learners
- Alternative Hypothesis: At least one LMS variant differs significantly from the others in UEQ or SUS

```
Friedman Tests:
Friedman Test (SUS): χ²=3.805, p=0.1492
Friedman Test (UEQ): χ²=1.238, p=0.5385
```

Figure 5.42: Friedman Test

The figure 5.42, displays the p-values from Friedman tests on both SUS and UEQ scores. For SUS scores it indicates p-value = 0.149. Although it is a low p value, but it is higher than the marginal significance ($p > 0.05$). This non-significant result ($p > 0.05$) indicates we cannot reject the null hypothesis that all three LMS variants have equal usability distributions when considering the entire sample simultaneously. For UEQ it displays p-value = 0.538, since 0.538 > 0.05, we fail to reject the null hypothesis, suggesting that we find no evidence of differences in UEQ scores across the LMS variants.

Despite the non-significant Friedman test results, we proceeded with post-hoc Wilcoxon signed-rank tests [26] to further compare the three LMS because Friedman test sometimes may miss the real differences, especially with small samples, so additional checks were needed.

5.5.5 Post-hoc Analysis

To identify which specific pairs of LMSs differ significantly, pairwise comparisons were conducted using the Wilcoxon Signed-Rank Test [26]. Bonferroni correction was applied to control for Type I error due to multiple comparisons.

Final Results with Adjusted p-values:

SUS Scores:

Comparison	W	p	d	adj_p
Base vs Manual	69.5	0.0634193	-0.57554	0.190258
Base vs AI	38.5	0.00285149	-0.8084	0.00855446
Manual vs AI	33.5	0.0734143	-0.403083	0.220243

UEQ Scores:

Comparison	W	p	d	adj_p
Base vs Manual	120	0.848594	-0.0508071	1
Base vs AI	72.5	0.22497	-0.274302	0.674911
Manual vs AI	98.5	0.808234	-0.245433	1

Figure 5.43: Post-hoc Analysis (Wilcoxon Signed-Rank Tests with Bonferroni Correction)

The figure 5.43 displays the results of each pairs of LMS with adjusted p values. According to above results we can identify that SUS score, Base LMS vs. AI-Generated LMS pair indicates p = 0.0085, representing a statistical significant difference ($p < 0.0167$). All the other pairs are not performing statistically significant difference ($p < 0.0167$). In order to this results we can ensure that Base vs AI comparison reached statistical significance with large effect size (d = -0.81) confirming the practical importance.

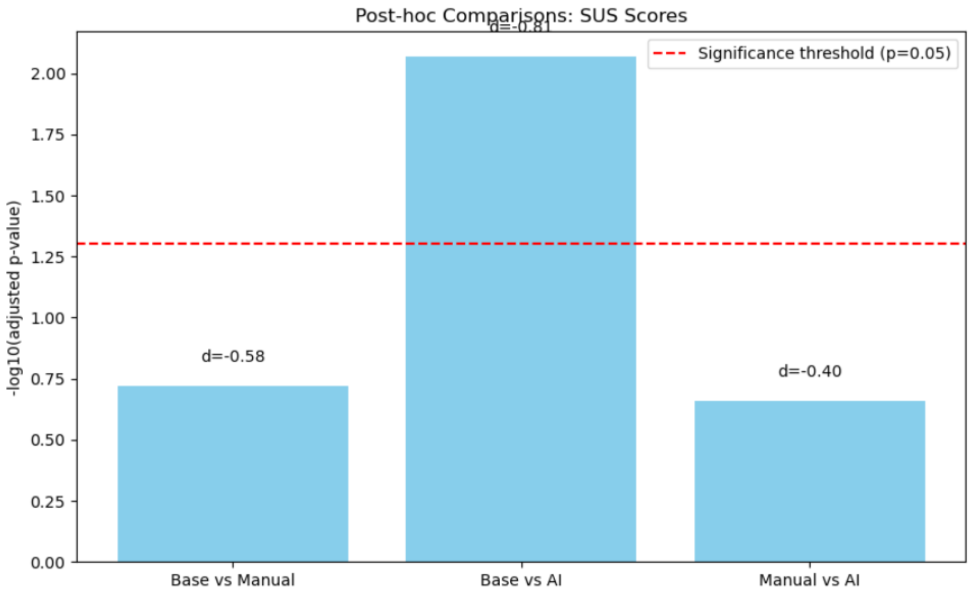


Figure 5.44: Post-hoc comparison of SUS scores

The figure 5.44, displays the post-hoc comparisons of SUS scores. It indicates the pair of Base LMS vs AI performs a statistically significant difference, with a large effect size.

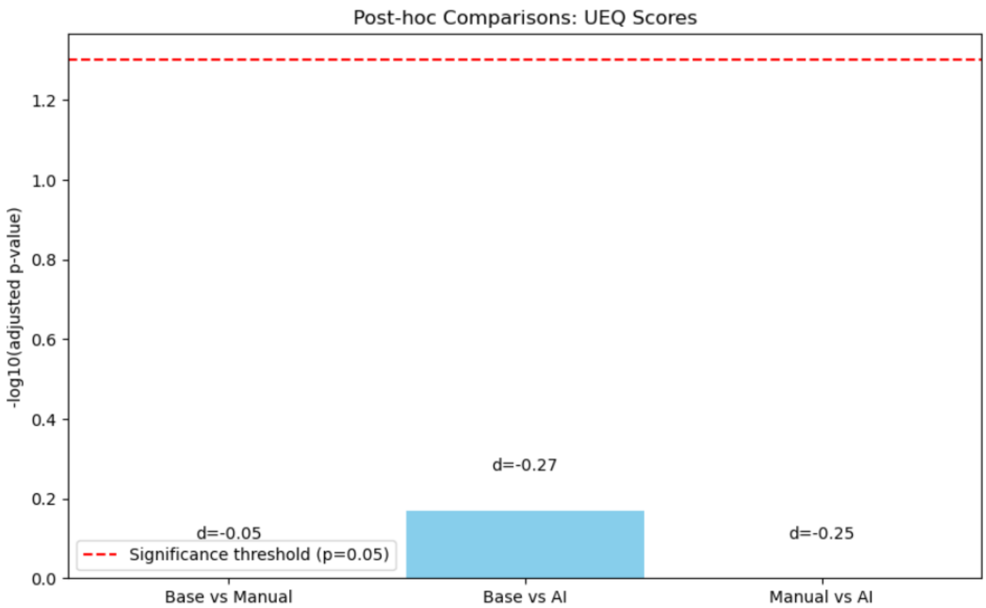


Figure 5.45: Post-hoc comparison of UEQ scores

The figure 5.45, displays the post-hoc comparisons of UEQ scores. There we cannot identify any significantly different pairs.

These results imply that AI LMS is the only system which shows statistically better usability than conventional LMS. Manual optimizations alone may not be sufficient for significant ADHD usability improvements. As well as it implies that while AI improves usability with SUS, core UX dimensions with UEQ still remain similar across systems. It indicates that UEQ may lack sensitivity to ADHD-specific UX improvements, suggesting the need for more specialized UX metrics.

5.5.6 Correlation Analysis

We conducted Spearman's rank correlation analysis [12] to examine the relationship between System Usability Scale (SUS) and User Experience Questionnaire (UEQ) scores across three LMS[12].

```
SUS-UEQ Correlations:
Base: ρ=0.373, p=0.0870
Manual: ρ=0.263, p=0.2365
AI: ρ=-0.158, p=0.4836
```

Figure 5.46: Correlation Analysis- Spearman's between SUS and UEQ

According to the figure 5.46, Base LMS shows $\rho = 0.373$ (p = 0.087) indicating a moderate positive relationship. It fails to reach significance ($p < 0.005$) may due to high variability in ADHD responses and limited sample size (n=22). Manually improved LMS shows $\rho = 0.263$(p = 0.237), indicating a weak positive relationship (not significant). AI LMS represents $\rho = -0.158$(p = 0.484), indicating a negligible inverse relationship (not significant). According to the above figures,there was no significant correlations found between SUS and UEQ scores.

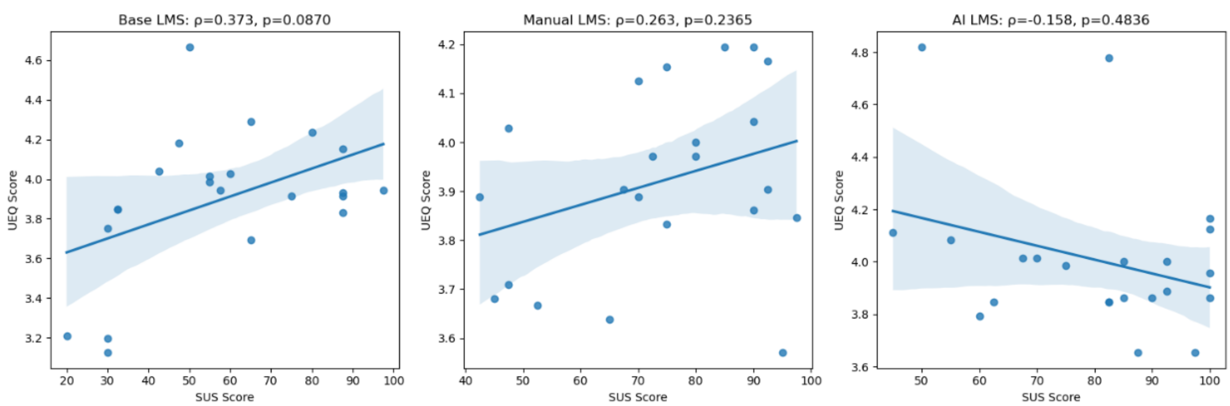


Figure 5.47: Scatter Plots of SUS vs UEQ

The figure 5.47 displays the scatterplots of UEQ and SUS for each LMS. Base LMS and Manual LMS show a moderate positive trend while AI LMS shows a negative relationship.

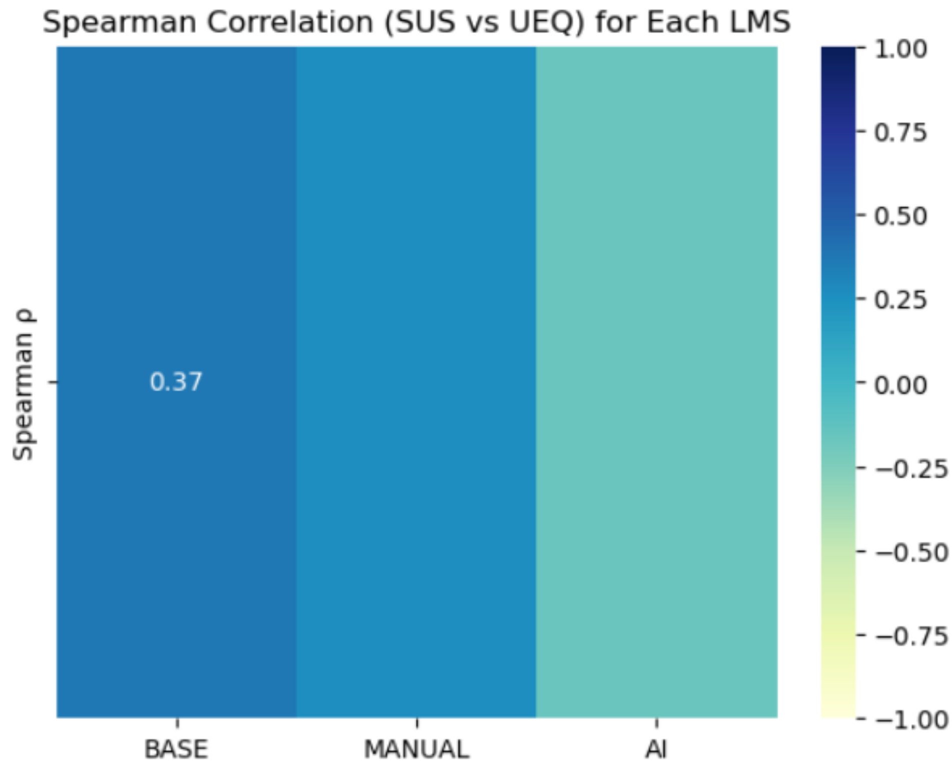


Figure 5.48: Spearsman Correlation

The figure 5.48 represents the spearman correlation variance across the three LMS.

5.5.7 Analysis of LMS Evaluation Scores Across ADHD Levels

To perform a better analysis of how the ADHD levels of the users have affected for the evaluations of each prototype, we compared the three LMSs — Base, Manual, and AI LMS based on user feedback from participants across different ADHD levels using usability and effectiveness metrics including SUS, UEQ, and Quiz scores. As indicated in the below figure 5.49 average SUS, UEQ and Quiz scores of each LMS were compared across the ADHD levels of the participants.

	ADHD Level	Base SUS Average	Manual SUS Average	AI SUS Average	Base UEQ Average	Manual UEQ Average	AI UEQ Average	Base Quiz Average	Manual Quiz Average	AI Quiz Average
0	High negative	52.50	87.00	91.50	22.85	24.12	23.78	5.00	6.20	5.20
1	Low negative	77.50	70.00	73.75	23.67	23.80	23.92	6.00	7.50	7.50
2	Low positive	57.17	69.83	77.17	23.53	23.28	24.15	5.47	6.53	5.87

Figure 5.49: Analysis of LMS Evaluation Scores Across ADHD Levels

Evaluation of SUS Scores Across ADHD Levels

When considering the average SUS scores acquired by different ADHD level participants across each of the LMS, as depicted in the figure 5.50, for Base LMS low negative individuals have acquired highest average SUS score compared to other two prototypes. The low positive individuals have performed slightly more preference for Base LMS compared to the high negative individuals. It indicates that Base LMS is comparatively more preferred by the individuals with lowest level of ADHD symptoms representing it is better for the general population. When considering the Manual LMS, high negative individuals have performed more preference compared two other ADHD levels. It represents that Manual LMS has qualified for facilitating the unique needs of the ADHD individuals because high negative level indicates the individuals with more ADHD symptoms. AI LMS is also more preferred by high negative individuals with the highest averge SUS score, indicating that LMS has become more usable for the ADHD students.

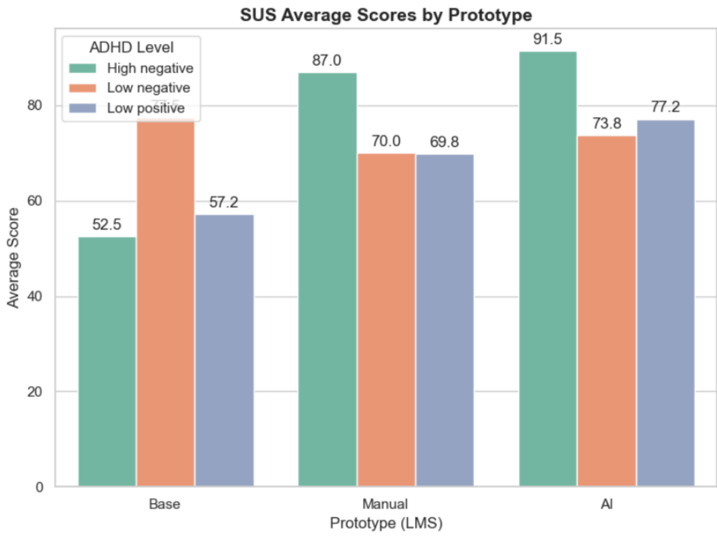


Figure 5.50: Analysis of SUS Scores Across ADHD Levels

Evaluation of UEQ Scores Across ADHD Levels

According to the figure 5.51, we compared the average UEQ scores variation across each of the ADHD level. When considering these UEQ results, we identified there was no a significant difference between the average UEQ scores acquired by each ADHD levels. For Base LMS similar to the SUS score results, here also majority of low negative individuals have rated, representing Base LMS is more suitable for the general population. For Manual and AI LMS high negative individuals with more ADHD symptoms, shows highest preference than the other two ADHD level categories, indicating these two prototypes have customized for ADHD population.

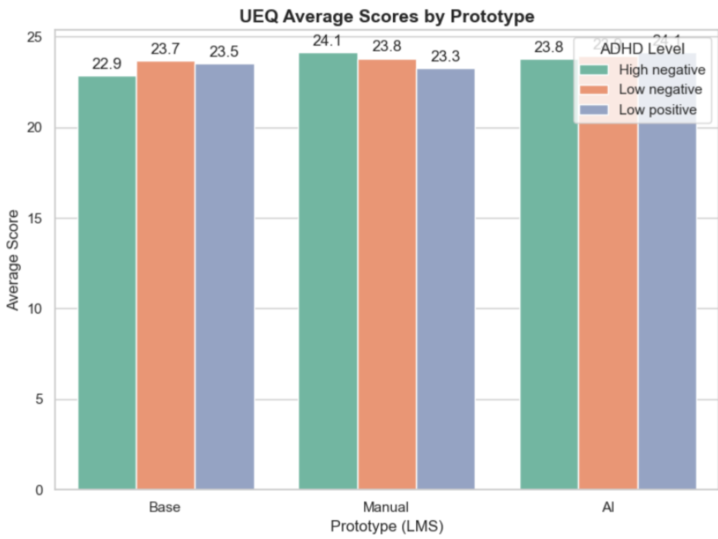


Figure 5.51: Analysis of UEQ Scores Across ADHD Levels

Evaluation of Quiz Scores Across ADHD Levels

Figure 5.52 shows the comparison of the average Quiz scores of different ADHD levels variation across each of the LMS. Overall it represents most of the low negative individuals have acquired highest scores across all of the three LMS. Then low positive individuals have scored comparatively more than high negative individuals across all the LMS. It implies that quiz performance has not been much positively supported for ADHD population.

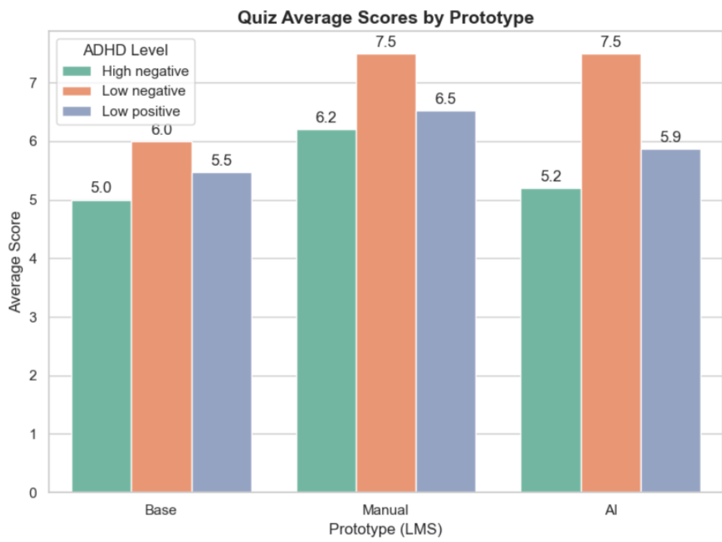


Figure 5.52: Analysis of Quiz Scores Across ADHD Levels

5.5.8 Microsoft Clarity Analysis

Prototype 2 achieved the highest overall performance score (82/100), followed by Prototype 3 (80/100) and Prototype 1 (77/100) (See Figure 5.53). While these differences appear modest, the performance breakdown revealed more details. Prototype 3 achieved the highest percentage of "good" ratings (61.9%) despite having the only "poor" ratings (3.2%). Largest Contentful Paint (LCP) metrics revealed significant loading performance differences, with Prototype 2 achieving the fastest content rendering (4.5s), substantially outperforming Prototype 3 (6.5s) and Prototype 1 (7.6s). This faster visual completion likely contributed to Prototype 2's higher engagement metrics by reducing initial waiting time. Interaction to Next Paint (INP) scores show both Prototype 2 (140ms) and Prototype 3 (150ms) delivering responsive interactions, while Prototype 1 lagged significantly (250ms). The Cumulative Layout Shift (CLS) metrics show all prototypes maintaining good visual stability during loading, with Prototype 2 achieving perfect stability (0). These technical metrics suggest that Prototype 2 delivered the most optimized technical performance, likely contributing to its superior user engagement patterns by providing a more responsive and consistent experience.

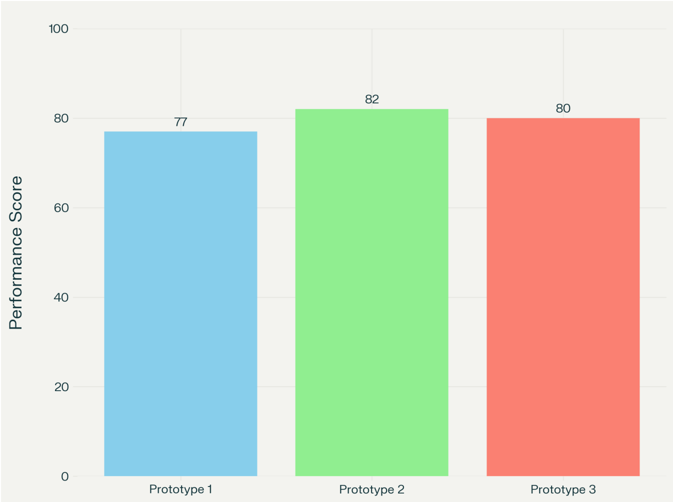


Figure 5.53: Overall Performance Scores

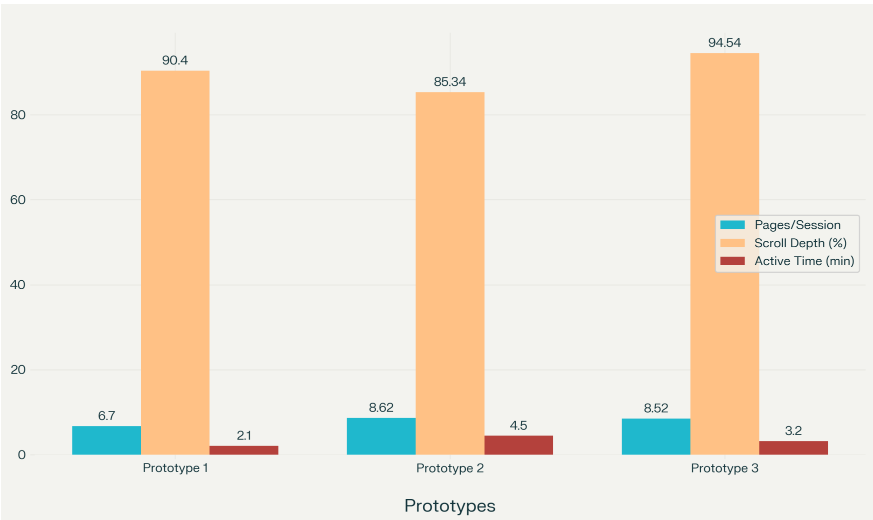


Figure 5.54: Engagement Metrics Comparison

Analysis of engagement metrics reveals that Prototypes 2 and 3 outperformed Prototype 1 in user exploration, with higher pages per session (8.62 and 8.52 vs. 6.7), indicating more comprehensive navigation. Prototype 3 achieved the greatest scroll depth (94.54%), suggesting users viewed more content per page, while Prototype 2 led in active engagement time (4.5 minutes), more than doubling that of Prototype 1. Overall, these results demonstrate that design enhancements in Prototypes 2 and 3 successfully promoted both deeper content exploration and sustained user interaction compared to the baseline. (See Figure 5.54)

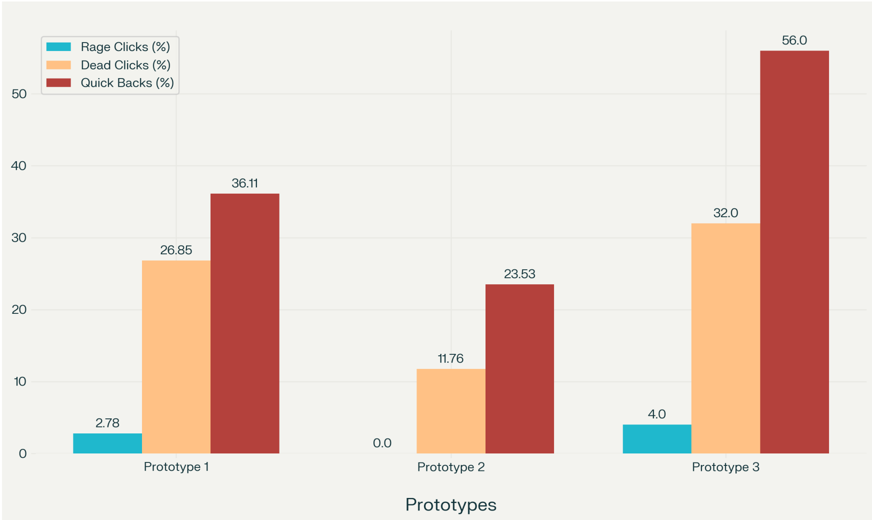


Figure 5.55: Interaction Metrics Comparison

The interaction metrics comparison revealed Prototype 2's better usability across all friction indicators, with 0% rage clicks, 11.76% dead clicks, and 23.53% quick backs. It indicated an interface where elements behaved as expected and navigation paths met user expectations. In contrast, Prototype 3 exhibited the highest user friction across all metrics (4% rage clicks, 32% dead clicks, and an alarming 56% quick backs), indicating some usability problems where more than half of page navigations resulted in immediate user dissatisfaction. Prototype 1 showed intermediate performance with moderate friction indicators (2.78% rage clicks, 26.85% dead clicks, and 36.11% quick backs). These patterns suggest that while Prototype 3 achieved high content exposure as shown in engagement metrics, its interface design created significant confusion through misleading visual affordances and poor information architecture. (See Figure 5.55)

In conclusion, Prototype 2 emerged as the most effective learning management system, combining the highest overall performance score and technical optimization with the best user engagement and lowest friction, as evidenced by its leading active time, minimal rage and dead clicks, and lowest quick back rate. Prototype 3, while achieving the highest content exposure through scroll depth and strong navigation, had issues with user frustration and navigation dissatisfaction. Prototype 1, which served as the baseline, showed moderate performance and engagement, but was outperformed by both redesigned systems. Overall, the results highlighted that balanced design improvements in Prototype 2 led to both enhanced learning engagement and a smoother, more effective user experience.

Chapter 6

Discussion and Conclusion

6.1 Discussion

This research aimed to determine how GenAI-generated customisation can increase the effectiveness and usability of LMSs for students with ADHD when compared to traditional and manually optimised platforms. We explored the impact of design and technology interventions on user effectiveness, usability, engagement and learning outcomes by creating and testing our three different prototypes.

Our findings directly addressed the primary research question, **How can AI-driven customization improve the e!ectiveness and usability of LMS for individuals with ADHD compared to conventional and manually optimized LMS platforms?**.
The GenAI generated LMS showed advantages in usability compared to the Base LMS and the Manual LMS. The participants comments such as "This is much better user friendly", and "Much better UI" confirmed the quantitative analysis results indicating that the usability of the AI LMS is better. On the other hand the performance of the Manual LMS showed slightly more advantages in effectiveness compared to the GenAI generated LMS. The participants comments like "I liked the new animations and stuff. They were engaging and kept my attention. And the focus mode function was helpful.", "The graphics and animations that you used were amazing. And specially the focus mode was helpful.", and "The study timer was a good addition. " showed that the effectiveness of the system in Manual LMS is higher than the AI LMS agreeing with the quantitative data received. Meanwhile some comments like "It's nice. but too distracting." was also received.

In addressing the first and second secondary questions, **How can user data be effectively used to evaluate the usability of conventional LMS to identify limitations and areas for improvement? What improvements are necessary in personalized learning systems to support ADHD students?**
our analysis of user data from the conventional LMS prototype revealed key usability limitations. We received comments such as "Absent of the back button made me irritated. I don't like to use browser's back button on websites.", "No back button in the system.",

and “Please show which were the right and wrong answers in quiz. “The participants comments mainly raised these issues. These insights influenced specific improvements in the prototypes highlighting the importance of user data. The Manual LMS was integrated with ADHD-friendly features like focus mode, engaging animations, and study timers. While these improvements were well-received by most users some participants found the system a bit more complex than needed, which were mentioned in their comments - “It’s nice. but too distracting.” .

The third secondary question, **How can prompting strategies in generative AI be optimized to create user interfaces that enhance engagement and usability for ADHD learners?**
explored how generative AI prompting strategies can be optimized to enhance engagement and usability. By using prompt engineering techniques, we were able to generate contextually relevant interfaces, resulting in more engaging and accessible interfaces for students with ADHD. Participants appreciated features such as progressive content disclosure and interactive quiz reviews. We received comments like “Question review after quiz was good feature”, “The way the content was presented was good, not showing everything at once.”, and “I do prefer if some interactive elements and nice to see things were there”. These findings highlighted both the opportunities and challenges of using GenAI in educational contexts, which could be addressed with better and more effective prompt engineering techniques.

Addressing the final secondary research question, **How can the impact of GenAI-driven personalized learning interfaces on academic performance and usability for individuals with ADHD be assessed?**
these findings highlighted both the opportunities and challenges of using GenAI in educational contexts, including the need for robust prompt engineering. Even though the usability of the AI LMS was better, in the educational context it showed that the manual LMS performed better. The Manual LMS achieved a slightly better quiz scores than the AI LMS. But the use of more advancedd Fenerative AI techniques could improve these results, leading to better effectiveness in educational context.

6.2 Conclusion

The objective of this research was to explore the importance of developing enhanced personalized learning interfaces using Generative AI for individuals with ADHD, who are students or undergraduates between the age range of 18 to 30, to improve their learning experience creating an equitable opportunity for them to excel in academics. Through the comprehensive literature review, surveys and user feedback we identified the unique challenges faced by the ADHD students in their conventional educational environments. For addressing their specific learning requirements, our research involved evaluating and comparing the usability and effectiveness of three distinct Learning Management Systems (LMS): a conventional LMS (Base LMS) used by the general population, a manually optimized LMS designed specifically for ADHD users, and an AI-generated LMS customized to ADHD needs through enhanced prompting techniques. Through this study our aim was to identify the most suitable learning system in terms of usability and effectiveness.

To ensure both usability and effectiveness of the system, we collected SUS, UEQ responses from our user group for measuring usability and used quiz scores analysis assessing learning outcomes and Microsoft clarity results for measuring the effectiveness. Based on the the SUS, UEQ, Quiz score results we conducted advanced statistical analysis with different statistical methods for deriving more reliable results.

As a conclusion, when comparing the usability of the prototypes considering SUS scores, AI LMS consistently received the highest SUS scores with highest mean and lowest standard deviation, indicating excellent usability. The Manual LMS scored moderately, while the Base LMS performed poorly. According to the UEQ score comparison , we identified that AI LMS scored the highest across all UEQ dimensions, especially in novelty supporting for maintaining focus and engagement among ADHD learners. The Manual LMS showed moderate improvement over the Base LMS, but still perform poor in areas like novelty. By achieving the highest performance in both SUS and UEQ scores analysis, AI Generated LMS performs as the optimal learning system in considering the usability of the system.

When evaluating the effectiveness of the prototypes by using quiz scores for measuring learning outcomes performed by the participants, Manual LMS acquired significantly highest mean quiz scores compared to the other two prototypes. AI LMS performed slightly higher quiz scores compared to the Base LMS. Overall it represents that Manual LMS performs better in effectiveness of the system. Further Microsoft clarity results also ensures that Manual LMS performs well for ADHD users in terms of system effectiveness.

According to the Normality test results we identified that most of the p values indicated a normal distribution while some of the values were non-normal. Based on that we further conducted analysis with Friedman non-parametric statistical test and Post hoc analysis using Wilcoxon Signed-Rank Test. According to these results we identified a statistical significance difference between Base LMS and AI LMS, implying that AI LMS shows statistically better usability than conventional LMS. Based on Spearman's Correlation analysis, we found there was no significant correlations between the SUS and UEQ scores across all LMS types, implying these two metrics may independently affecting for the system usability.

By analysing the LMS Evaluation Scores Across ADHD Levels, we found that when analysing SUS and UEQ scores, Base LMS is more preferred by the low negative level ADHD users, representing that prototype is more applicable for the general population. Manul LMS and AI LMS were more preferred by high negative ADHD individuals with more ADHD symptoms, representing these two prototypes are performing better for ADHD population compared to the Base LMS. It implies that customized AI and Manual LMS were more usable and effective for the ADHD students for enahncing their learning experience.

Considering evaluation metrics, AI LMS can be identified as the best-fitting prototype for the ADHD students in the usability aspect, while Manual LMS appeared better in the dimension of effectiveness. Therefore both AI-enhanced LMS and manually optimized LMS plays a significant role in enhancing the overall learning experience of ADHD students addressing their unique educational needs and facilitating their unique behavioural and cognitive requirements. To the contrary, the base LMS continues to be more appropriate for the larger student population that does not have specific learning requirements due to its generalised design.

Chapter 7

Recommendations and Future Work

7.1 Recommendation

During the conduct of this research, creating the AI LMS was very challenging because of the limited capabilities of LLMs to understand the system architecture of a web application. This limitation resulted in architectural inconsistencies, error propagation, and decreased system reliability. Development of the AI LMS required significant human oversight and architectural guidance.

Future research directions focusing on reference architectures, iterative approaches, and architectural guidance systems offer potential pathways to address these limitations. It is recommended to use an AI Model with system architectural awareness in the future research.

7.2 Future Work

Using these research findings, a variety of new possibilities can be explored to extend the development of personalized learning interfaces for ADHD students. Incorporating eye tracking in evaluations to identify real time patterns of visual focus and identify interface elements which enhances or diminishes attention. This method could also be combined with clickstream analysis, to link design features (e.g., color schemes, interactive queries) with task-completion rates and enable more data-informed improvements. Furthermore, longitudinal studies could be used to explore the impact of these Manual LMS and AI LMS usage on long-term knowledge retention and better academic performance. Usage of wearable biometric sensors like EEG headsets, and heart-rate monitors could be used to examine deeper physiological responses to interface adjustments. Real-time biometric feedback loops such as galvanic skin response sensors could be used to automatically adjust attention retention methods. Moreover, methods could be explored to adapt and personalize the interfaces for each individual ADHD student for their specific ADHD traits by analyzing behavioural patterns in real time to better suit individual neurocognitive profiles. Multimodal options could be explored and developed so that the interfaces dynamically evolve with the users, enabling success in academics.

Additionally, future studies might gain from the creation and application of specialized AI models specifically designed to assist individuals with ADHD or comparable conditions. In contrast to this research, which used prompting strategies to modify general-purpose models, a tailored model could provide a more individualized and effective user experience suited to cognitive and behavioral needs.

In the future, it's important to carry out additional research with a larger group of students who are medically diagnosed with ADHD to acquire a deeper insight into their academic difficulties, cognitive behaviors, and the efficacy of customized educational interventions.

References

[1] Tomás Alves, Joana Natálio, Joana Henriques-Calado, and Sandra Gama. Incorporating personality in user interface design: A review. *Personality and Individual Differences*, 151:109709, 2019. doi: 10.1016/j.paid.2019.109709.

[2] Mental Health America. Adhd screening tool, 2023. URL `https://screening.mhanational.org/screening-tools/adhd/`.

[3] Giovanni Anobile, Massimo Bartoli, Gabriele Masi, Andrea Tacchi, and Francesca Tinelli. Math difficulties in attention deficit hyperactivity disorder do not originate from the visual number sense. *Frontiers in Human Neuroscience*, 16:949391, 2022. doi: 10.3389/fnhum.2022.949391.

[4] Bagustari A. Bagustari and Hery B. Santoso. Adaptive user interface of learning management systems for education 4.0: A research perspective. *Journal of Physics: Conference Series*, 1235(1):012033, 2019. doi: 10.1088/1742-6596/1235/1/012033.

[5] Prabal Datta Barua, Jahmunah Vicnesh, Raj Gururajan, Shu Lih Oh, Elizabeth Palmer, Muhammad Mokhzaini Azizan, Nahrizul Adib Kadri, and U. Rajendra Acharya. Artificial intelligence enabled personalised assistive tools to enhance education of children with neurodevelopmental disorders—a review. *International Journal of Environmental Research and Public Health*, 19(3):1192, 2022. doi: 10.3390/ijerph19031192. URL `https://www.mdpi.com/1660-4601/19/3/1192`.

[6] Richard Baskerville, Abayomi Baiyere, Shirley Gregor, Alan Hevner, and Matti Rossi. Design science research contributions: Finding a balance between artifact and theory. *Journal of the Association for Information Systems*, 19:358–376, 05 2018. doi: 10.17705/1jais.00495.

[7] Santiago Berrezueta-Guzman, Mohanad Kandil, María-Luisa Martín-Ruiz, Iván Pau de la Cruz, and Stephan Krusche. Future of adhd care: Evaluating the efficacy of chatgpt in therapy enhancement. *Healthcare*, 12(6):683, 2024. doi: 10.3390/healthcare12060683. URL `https://www.mdpi.com/2227-9032/12/6/683`.

[8] Santiago Berrezueta-Guzman, Mohanad Kandil, María-Luisa Martín-Ruiz, Iván Pau de la Cruz, and Stephan Krusche. Future of adhd care: Evaluating the efficacy of chatgpt in therapy enhancement. *Healthcare*, 12(6):683, June 2024. doi: 10.3390/healthcare12060683.

[9] John Brooke. Sus: A 'quick and dirty' usability scale. *Usability Evaluation in Industry*, pages 189–194, 1996. `https://www.researchgate.net/publication/228593520_SUS_A_quick_and_dirty_usability_scale`, Accessed: 2025-04-17.

[10] Alessio Bucaioni, Martin Weyssow, Junda He, Yunbo Lyu, and David Lo. A functional software reference architecture for llm-integrated systems, 2025. URL `https://arxiv.org/abs/2501.12904`.

[11] G. Barrera Castro, A. Chiappe, D. F. Becerra Rodriguez, and F. G. Sepulveda. Harnessing ai for education 4.0: Drivers of personalized learning. *Electronic Journal of e-Learning*, 22(5):1–16, 2024. doi: 10.34190/ejel.22.5.3467. URL `https://doi.org/10.34190/ejel.22.5.3467`.

[12] DATAtab Team. Spearman's rank correlation coefficient: A beginner's guide. `https://datatab.net/tutorial/spearman-correlation`, 2025. Accessed: 2025-04-17.

[13] DATAtab Team. Friedman test tutorial: Analyzing repeated measures data. `https://datatab.net/tutorial/friedman-test`, 2025. Accessed: 2025-04-17.

[14] Holly E. Erskine, Rosana E. Norman, Alize J. Ferrari, Gary C. K. Chan, William E. Copeland, Harvey A. Whiteford, and James G. Scott. Long-term outcomes of attention-deficit/hyperactivity disorder and conduct disorder: A systematic review and meta-analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*, 55(10):841–850, 2016. doi: 10.1016/j.jaac.2016.06.016. URL `https://pubmed.ncbi.nlm.nih.gov/27663939/`.

[15] Ryan Ferguson, Paul W. Cuddihy, Ed Price, Jacob Morin, John A. Buder, and Vincent J. Berardi. Further development of the f.o.c.u.s.: Facilitating on-going concentration using stimulation. In *2013 IEEE 39th Annual Northeast Bioengineering Conference*, pages 209–210, 2013. doi: 10.1109/NEBEC.2013.6574469. URL `https://ieeexplore.ieee.org/document/6574469`.

[16] Lydia Furman. What is attention-deficit hyperactivity disorder (adhd)? *Journal of Child Neurology*, 20(12):994–1002, 2005. doi: 10.1177/08830738050200121301.

[17] Miranda Gaub and Caryn L. Carlson. Gender differences in adhd: A meta-analysis and critical review. *Journal of the American Academy of Child & Adolescent Psychiatry*, 36(8):1036–1045, 1997. doi: 10.1097/00004583-199708000-00011. URL `https://www.sciencedirect.com/science/article/pii/S0890856709626125`.

[18] Vasiliki Gkora. Advancing adhd education: autonomy, technology, and inclusive strategies. *GSC Advanced Research and Reviews*, 18(3):101–111, March 2024. doi: 10.30574/gscarr.2024.18.3.0084.

[19] Atefeh Mahdavi Goloujeh, Anne Sullivan, and Brian Magerko. Is it ai or is it me? understanding users' prompt journey with text-to-image generative ai tools. In *CHI '24: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, pages 1–13. ACM, 2024. doi: 10.1145/3613904.3642861. URL https://doi.org/10.1145/3613904.3642861.

[20] Nielsen Norman Group. 5 visual treatments that improve accessibility. https://www.nngroup.com/articles/visual-treatments-accessibility/, 2022. Accessed: 2025-04-17.

[21] Jan Gulliksen, Bengt Göransson, Inger Boivie, Stefan Blomkvist, Jenny Persson, and Åsa Cajander. Key principles for user-centred systems design. *Behaviour & Information Technology*, 22(6):397–409, 2003. doi: 10.1080/01449290310001624329. URL https://doi.org/10.1080/01449290310001624329.

[22] M. Ibrahim, P. W. C. Prasad, Abeer Alsadoon, and L. Pham. Synchronous virtual classroom for students with adhd disorder. In *2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, pages 1–6, 2016. doi: 10.1109/JCSSE.2016.7748860. URL https://doi.org/10.1109/JCSSE.2016.7748860.

[23] Roberto Henriques Vladimir Trajkovik Ivica Pesovski, Ricardo Santos. Generative ai for customizable learning experiences. *Sustainability*, 16(7):3034, July 2024. doi: 10.3390/su160703034. Available online.

[24] Vibha Kaushik and Christine A. Walsh. Pragmatism as a research paradigm and its implications for social work research. *Social Sciences*, 8(9), 2019. ISSN 2076-0760. doi: 10.3390/socsci8090255. URL https://www.mdpi.com/2076-0760/8/9/255.

[25] Ronald C. Kessler, Thomas A. Adler, Michael J. Barkley, Daniel M. Biederman, John T. Faraone, Stephen V. Spencer, and David M. Zaslavsky. The world health organization adult adhd self-report scale (asrs): A short screening scale for use in the general population. *Psychological Medicine*, 35(2):245–256, 2005. doi: 10.1017/S0033291704002892. URL https://pubmed.ncbi.nlm.nih.gov/15841682/. Accessed: 2023-01-17.

[26] Laerd Statistics. Wilcoxon signed-rank test using spss statistics. https://statistics.laerd.com/spss-tutorials/wilcoxon-signed-rank-test-using-spss-statistics.php, 2025. Accessed: 2025-04-17.

[27] Vivian Liu and Lydia B. Chilton. Design guidelines for prompt engineering text-to-image generative models. In *CHI '22: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–23. ACM, 2022. doi: 10.1145/3491102.3501825. URL https://doi.org/10.1145/3491102.3501825.

[28] Evanfiya Logacheva, Arto Hellas, James Prather, Sami Sarsa, and Juho Leinonen. Evaluating contextually personalized programming exercises created with generative ai. *arXiv*, abs/2407.11994, 2024. URL https://arxiv.org/abs/2407.11994.

[29] Gianluca Malato. An introduction to the shapiro-wilk test for normality. https://builtin.com/data-science/shapiro-wilk-test, 2025. Accessed: 2025-04-17.

[30] Laura Mancera, Silvia Baldiris, Ramón Fabregat, Sergio Gomez, and Carolina Mejia. atenderah: A videogame to support e-learning students with adhd. In *2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT)*, pages 438–440, 2017. doi: 10.1109/ICALT.2017.157.

[31] Lorna McKnight. Designing for adhd: in search of guidelines. In *IDC 2010: Digital Technologies and Marginalized Youth (Interaction Design and Children)*, Barcelona, Spain, June 2010.

[32] G. Merlo, G. Chiazzese, M. Sanches-Ferreira, A. Chifari, L. Seta, C. McGee, A. Mirisola, and I. Giammusso. The whaam application: a tool to support the evidence-based practice in the functional behaviour assessment. *Journal of Innovation in Health Informatics*, 25(2):63–70, Jun 29 2018. doi: 10.14236/jhi.v25i2.919.

[33] Magdalena Mirkowicz and Grzegorz Grodner. Jakob nielsen's heuristics in selected elements of interface design of selected blogs. *Social Communication*, 4:30–51, 12 2018. doi: 10.2478/sc-2018-0013.

[34] Jakob Nielsen. Enhancing the explanatory power of usability heuristics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '94, pages 152–158, New York, NY, USA, 1994. Association for Computing Machinery. ISBN 0897916506. doi: 10.1145/191666.191729. URL https://doi.org/10.1145/191666.191729.

[35] Jakob Nielsen. Scenarios in discount usability engineering. In John M. Carroll, editor, *Scenario-Based Design: Envisioning Work and Technology in System Development*, pages 59–83. John Wiley & Sons, 1995. Accessed: 2025-04-17.

[36] Jakob Nielsen. 10 usability heuristics for user interface design. https://www.nngroup.com/articles/ten-usability-heuristics/, 2020. Accessed: 2025-04-17.

[37] Jakob Nielsen. Slips: User errors in user interfaces. https://www.nngroup.com/articles/slips/, 2023. Accessed: 2025-04-17.

[38] NovoPsych. Adult adhd self-report scale (asrs), 2023. URL https://novopsych.com.au/assessments/diagnosis/adult-adhd-self-report-scale-asrs/.

[39] NovoPsych. Adult adhd self-report scale (asrs), n.d. URL https://novopsych.com.au/assessments/diagnosis/adult-adhd-self-report-scale-asrs/. Accessed: February 18, 2025.

[40] R.C.W.M.R.H.E. Nugawela, W.G.S. Fonseka, G.U.C. Perera, S.K.S. Mihirangi, and K.A.D.C.P. Kahandawaarachchi. Therapy tool for adolescents with adhd. In *2020 IEEE 3rd International Conference on Communication, Computing and Digital Systems (C-CODE)*, pages 313–318, 2020. doi: 10.1109/C-CODE49350.2020.9103423. URL https://doi.org/10.1109/C-CODE49350.2020.9103423.

[41] Ken Peffers, Tuure Tuunanen, Marcus A. Rothenberger, and Samir Chatterjee. A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3):45–77, 2007. doi: 10.2753/MIS0742-1222240302. URL https://doi.org/10.2753/MIS0742-1222240302.

[42] Maria Perez-Ortiz, Claire Dormann, Yvonne Rogers, Sahan Bulathwela, Stefan Kreitmayer, Emine Yilmaz, Richard Noss, and John Shawe-Taylor. X5learn: A personalised learning companion at the intersection of ai and hci. In *Companion Proceedings of the 26th International Conference on Intelligent User Interfaces*, pages 70–74, 2021. doi: 10.1145/3397482.3450721. URL https://doi.org/10.1145/3397482.3450721.

[43] Ivica Pesovski, Ricardo Santos, Roberto Henriques, and Vladimir Trajkovik. Generative ai for customizable learning experiences. *Sustainability*, 16(7):3034, 2024. doi: 10.3390/su160703034. URL https://doi.org/10.3390/su160703034.

[44] Sandesh Phalke and Abhishek Shrivastava. *HI Applications for ADHD Children: The Graphic Design Perspective*, pages 105–127. Springer, 03 2024. ISBN 978-981-97-0231-2. doi: 10.1007/978-981-97-0232-9_2.

[45] Nina Rajcic, Maria Teresa Llano Rodriguez, and Jon McCormack. Towards a diffractive analysis of prompt-based generative ai. In *CHI '24: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, pages 1–15. ACM, 2024. doi: 10.1145/3613904.3641971. URL https://doi.org/10.1145/3613904.3641971.

[46] Shubham Rastogi, Anita Meena, R. S. Pandey, and R. K. Srivastava. Echo adhd in india: A feasible and acceptable training model for child-care physicians to identify and manage attention deficit hyperactivity disorder. *International Journal of Pediatric Research*, 7 (4):352–355, 2020. doi: 10.17511/ijcp.2020.i4.06. URL https://www.ijpediatrics.com/index.php/ijcp/article/view/5401.

[47] Vandana Ravishankar. Designing an assistive technology for self-reflection for students suffering from adhd at malmö university. Master's thesis, Malmö University, 2022. URL https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1701424&dswid=-8236.

[48] Martin Schrepp. User experience questionnaire handbook. https://www.researchgate.net/publication/281973617_User_Experience_Questionnaire_Handbook, 2015. Accessed: 2025-04-17.

[49] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. Applying the user experience questionnaire (ueq) in different evaluation scenarios. *Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience*, pages 383–392, 2014. doi: 10.1007/978-3-319-07668-3_37. https://www.researchgate.net/publication/265784386_Applying_the_User_Experience_Questionnaire_UEQ_in_Different_Evaluation_Scenarios, Accessed: 2025-04-17.

[50] Timothy K. Shih, Rong-Chi Chang, and Yi-Ming Chen. A personalized e-learning interface based on user's learning style. In *2007 European Conference on Computer Systems and Technologies (EURCON)*, pages 1–4, 2007. doi: 10.1109/EURCON.2007.4400362.

[51] Edmund J. S. Sonuga-Barke, Daniel Brandeis, Samuele Cortese, David Daley, Maite Ferrin, Martin Holtmann, Jim Stevenson, Argyris Stringaris, Eric Taylor, and Alessandro Zuddas. Nonpharmacological interventions for adhd: Systematic review and meta-analyses of randomized controlled trials of dietary and psychological treatments. *American Journal of Psychiatry*, 170(3):275–289, March 2013. doi: 10.1176/appi.ajp.2012.12070991.

[52] Safura Adeela Sukiman and Noraisyah Abdl Aziz. Artificial intelligence as the key player in learning interventions: A mini review among students with learning difficulties. *International Journal of Teaching and Learning in Higher Education*, 3(2):16450, 2021. URL https://myjms.mohe.gov.my/index.php/ijtmis/article/view/16450.

[53] UXmatters Editorial Team. Embracing neurodiversity in ux design: Crafting inclusive digital environments. https://www.uxmatters.com/mt/archives/2024/04/embracing-neurodiversity-in-ux-design-crafting-inclusive-digital-environments.php, 2024. Accessed: 2025-04-17.

[54] Maeghan L. Thomas and Judy L. Cezeaux. Further development of the f.o.c.u.s.: Facilitating on-going concentration in undergraduate students. In *2013 39th Annual Northeast Bioengineering Conference*, pages 283–284, 2013. doi: 10.1109/NEBEC.2013.19.

[55] Nathan Thomas. How to use the system usability scale (sus) to evaluate the usability of your website. https://usabilitygeek.com/how-to-use-the-system-usability-scale-sus-to-evaluate-the-usability-of-your-webs 2015. Accessed: 2025-04-17.

[56] Tien Tran, Hae-Na Lee, and Ji Hwan Park. Discovering accessible data visualizations for people with adhd. In *CHI '24: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, pages 1–19, 2024. doi: 10.1145/3613904.3642112. URL https://doi.org/10.1145/3613904.3642112.

[57] Jan vom Brocke, Alan Hevner, and Alexander Maedche. Introduction to design science research. In Jan vom Brocke, Alan Hevner, and Alexander Maedche, editors, *Design Science Research. Cases*, pages 1–13. Springer, Cham, 2020. doi: 10.1007/978-3-030-46781-4_1.

[58] Zhijie Wang, Zijie Zhou, Da Song, Yuheng Huang, Shengmai Chen, Lei Ma, and Tianyi Zhang. Towards understanding the characteristics of code generation errors made by large language models, 2025. URL https://arxiv.org/abs/2406.08731.

[59] Justin D. Weisz, Jessica He, Michael Muller, Gabriela Hoefer, Rachel Miles, and Werner Geyer. Design principles for generative ai applications. In *CHI '24: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, pages 1–22. ACM, 2024. doi: 10.1145/3613904.3642466. URL https://doi.org/10.1145/3613904.3642466.

[60] World Wide Web Consortium (W3C). Web content accessibility guidelines (wcag) 2.1, 2018. URL https://www.w3.org/TR/WCAG21/. W3C Recommendation, updated 2023-09-21. Published by the Accessibility Guidelines Working Group.

[61] Jason Wu, Eldon Schoop, Alan Leung, Titus Barik, Jeffrey P. Bigham, and Jeffrey Nichols. Uicoder: Finetuning large language models to generate user interface code through automated feedback, 2024. URL https://arxiv.org/abs/2406.07739.

[62] Han Yu, Chunyan Miao, Cyril Leung, and Timothy John White. Towards ai-powered personalization in mooc learning. *npj Science of Learning*, 2, 2017. doi: 10.1038/s41539-017-0016-3. URL https://www.nature.com/articles/s41539-017-0016-3.

Appendices

Appendix: A

- Primary Survey: https://forms.gle/4vCp68qNDM5tXekn9
- Primary GitHub Repository: https://github.com/RashmiNirasha/GenAIResearch.git
- User Flow and Tasks: https://docs.google.com/document/d/1AGKYcMceRiUCTHzNtKA9HM9ASlpGCl7aXPW7cGXZy0Q/edit?usp=sharing
- Heuristics and Laws Followed: https://docs.google.com/spreadsheets/d/1_QnjSndTRbhSYAjXkkPWHA-IV3w_EOlOSpEWxmSGAqQ/edit?gid=0#gid=0

Ethics Files

- Ethical Clearance Certificate https://drive.google.com/file/d/1VBub2wm9Q7KLYbelo8N6Ih0826ozcLDM/view?usp=sharing
- Information Sheet and Informed Consent Document https://drive.google.com/file/d/19HE7H95NBDgZvOrRSzUbzC5_wk3sDxKK/view?usp=sharing

Prototype 01: Base LMS

- GitHub repository: https://github.com/michellenikeetha/base_lms.git
- Deployed LMS: https://base-lms.vercel.app/
- Clarity Project: https://clarity.microsoft.com/projects/view/p1n97ondqi/settings#overview
- SUS Evaluation Form https://forms.gle/vtVgEvyiVAGH7FhW7
- UEQ Evaluation Form https://forms.gle/VxiEgV68e2xxQfyAA

Prototype 02: Manual LMS

- GitHub repository: https://github.com/michellenikeetha/manual_adhd_lms.git
- Deployed LMS: https://manual-adhd-lms.vercel.app/
- Clarity Project: https://clarity.microsoft.com/projects/view/qiw91emtrr/settings#overview
- SUS Evaluation Form https://forms.gle/XHdF7nEGUSqDcCH77
- UEQ Evaluation Form https://forms.gle/Lc9aroCB5Qs1iAjG6

Prototype 03: AI LMS

- GitHub repository: `https://github.com/RashmiNirasha/ai-vle.git`
- Deployed LMS: `https://ai-vle.vercel.app/login`
- Clarity Project: `https://clarity.microsoft.com/projects/view/qfsogio69t/settings?date=Last%203%20days#overview`
- SUS Evaluation Form `https://forms.gle/ovmmRRiMrVWX8QrA9`
- UEQ Evaluation Form `https://forms.gle/MQuysxWF9CgjhKUt7`

Generative AI Resources

- AI Interface Evaluation Survey `https://forms.gle/LzBJuogj7xuTUsXr5`
- ChatGPT Extension Link `https://chatgpt.com/g/g-67c5a45618f48191b44885e8b054241c-adhd-friendly-lms-engineer`
- Collection of Prompts Used `https://shorturl.at/QvWM4`

Evaluation: Statistical Analysis

- GitHub repository: `https://github.com/Rashi990/evalution_statistical_analysis.git`

Appendix: B

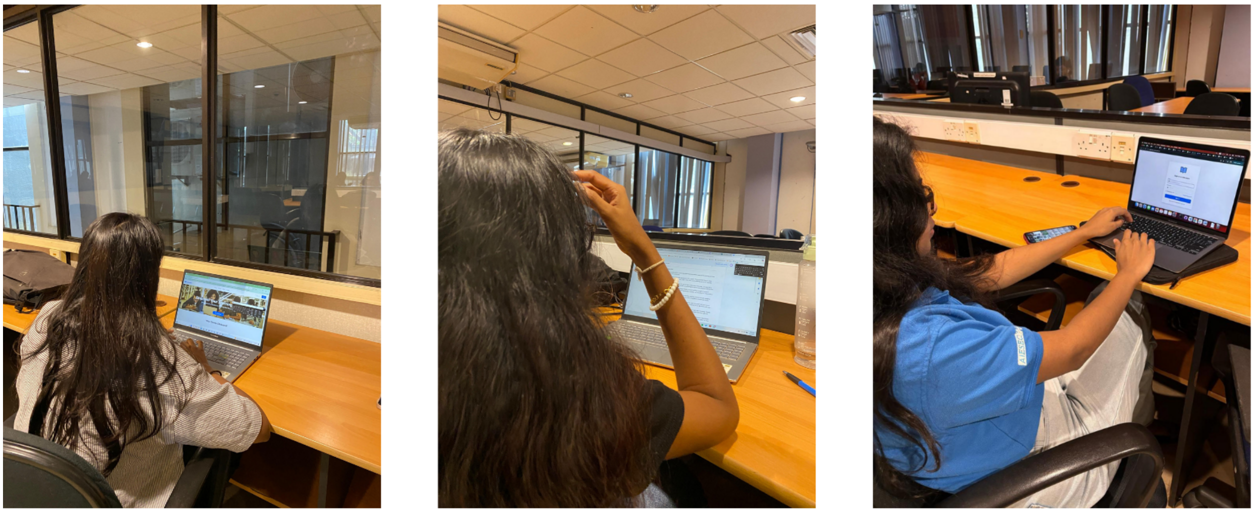


Figure 1: User Participation

MS Clarity Dashboards:

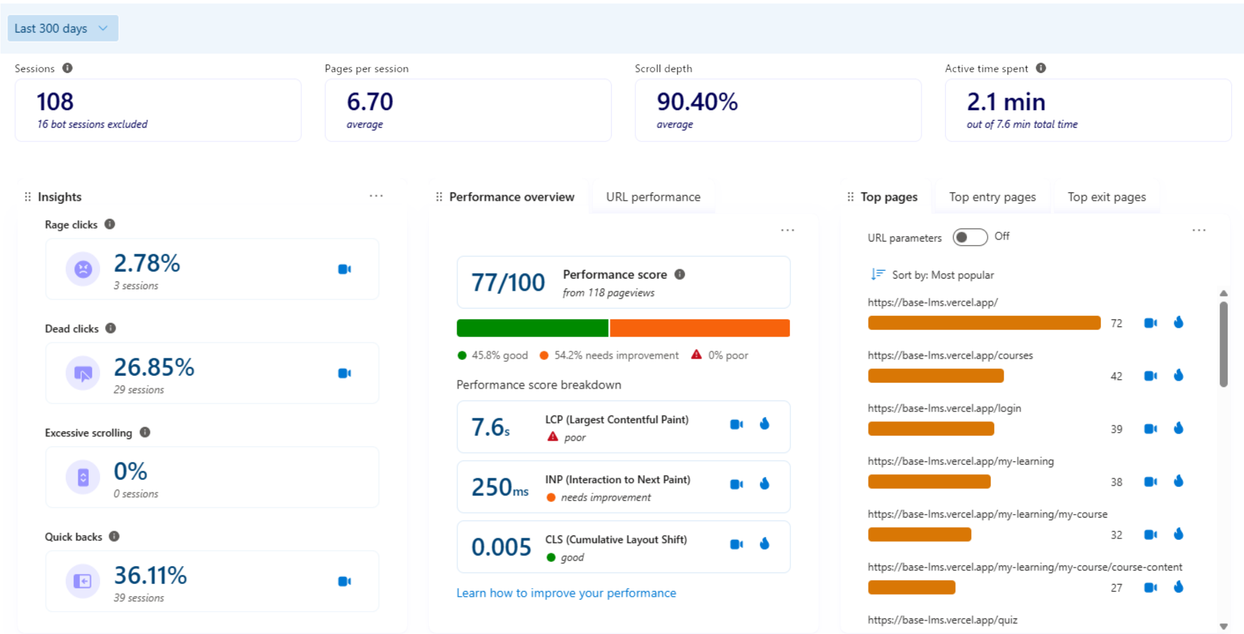


Figure 2: Base LMS: Clarity Dashboard

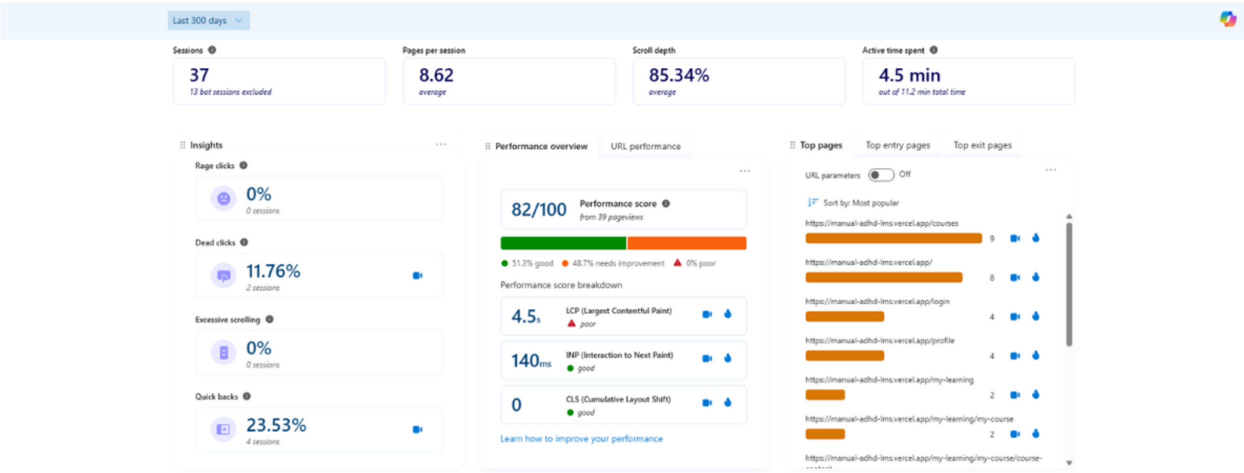


Figure 3: Manual LMS: Clarity Dashboard

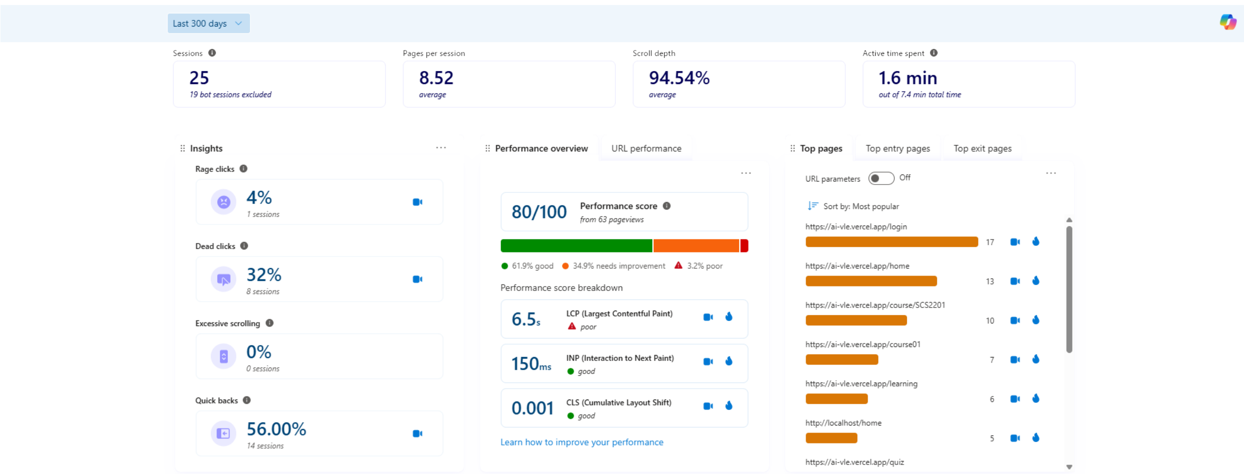


Figure 4: AI LMS: Clarity Dashboard