



FUSION 2025

Proceedings of the 7th National Symposium on Human Computer Interaction, 2025

“Humanizing Intelligence”

25 - 26 October 2025
Asia Pacific University of
Technology & Innovation
(APU)

Editors:

Siti Nurliana Jamalai@Jamali
Shaima Abdulfattah Gamal Mohammed
Samia Abdhood
Khairunnisa Binti Mispa

Organized by:



Supported by:



A P U
ASIA PACIFIC UNIVERSITY
OF TECHNOLOGY & INNOVATION



Chapter Conference

Fusion 2025

Proceedings of the 7th National Symposium on Human-Computer Interaction 2025

25- 26 October 2025

Asia Pacific University of
Technology & Innovation
(APU)

Editors:

Siti Nurliana Jamalai@Jamali

Shaima Abdulfattah Gamal Mohammed

Samia Abdhood

Khairunnisa Binti Mispa

Organized by:



Copyright ©2025.myHCI-UX

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Proceedings of the National Symposium on Human-Computer Interaction will be published annually by myHCI-UX.

eISSN: 3009-1624

Published by:
myHCI-UX
C/O MYHCI-UX, UTM IBS,
Level 10, Menara Razak,
Universiti Teknologi Malaysia,
Jalan Sultan Yahya Petra,
54100 Kuala Lumpur,
Malaysia.
contactus@myhci-ux.org

Welcome Messages

Foreword from Fusion2025 Chair

Dear Colleagues, Innovators, Researchers and Friends,



It is my great pleasure to welcome you to 7th National Symposium on Human-Computer Interaction (FUSION2025), a gathering of minds dedicated to shaping the future of interdisciplinary innovation. This year's theme, "Humanizing Intelligence", reflects our collective commitment to exploring the convergence of technology, design, and society. As Chair of FUSION2025, I am proud to see how this platform continues to evolve, bringing together researchers, practitioners, and students from diverse domains to share insights, challenge assumptions, and co-create new possibilities. From immersive technologies and AI ethics to inclusive design and digital transformation,

the conversations we foster here are more critical than ever.

This year's program features a rich tapestry of keynotes, panels, workshops, Doctoral consortium, paper presentation and competitions that highlight the power of collaboration across disciplines. I encourage you to engage deeply, ask bold questions, and connect with new perspectives. Whether you are presenting, participating, or simply exploring, your presence adds value to this vibrant community.

On behalf of the organizing committee, I thank you for being part of FUSION2025. May this experience inspire new ideas, meaningful partnerships, and lasting impact.

Warm regards,

Ts. Dr. Siti Azreena Binti Mubin

Chair, FUSION2025

Welcome Message

Foreword from myHCI-UX President



Welcome to the 7th National Symposium on Human-Computer Interaction 2025 (FUSION 2025)!

For the seventh consecutive year, this symposium brings together researchers, practitioners, educators and students working in the fields of Human-Computer Interaction (HCI) and User Experience (UX) across Malaysia. The aim of FUSION has always been to provide a dynamic, inclusive platform for knowledge- sharing, peer learning, and exploration of collaborative opportunities in HCI and UX design.

In addition to the paper presentations, FUSION 2025 also includes the Student Design Competition, which is designed to cultivate and highlight new talent in HCI and UX; and also the Doctoral Consortium, which serves as a supportive space for postgraduate researchers to enhance their studies, network with peers and mentors, and strengthen their research directions.

The importance of HCI and UX in the Malaysian context cannot be overstated. As Malaysia positions itself as the gateway to innovation within the ASEAN region, the digital economy is rapidly evolving. Malaysia's strategic role in ASEAN underscores the need for interactive technologies that are both usable and culturally attuned. At the same time, emerging technologies such as artificial intelligence (AI), extended reality (XR), the Internet of Things (IoT), and smart-home systems demand new interaction paradigms where human-centred design and usability are central.

FUSION 2025 therefore calls on participants to reflect not only on the design, evaluation and implementation of interactive systems, but also on the wider ecosystem, including issues of accessibility, cultural diversity, ethical technology, and both regional and global collaboration. By bringing together academic research, industry practice and student innovation under one roof, we aim to foster a Malaysian HCI-UX community that is attuned to both local needs and global standards.

This year's symposium has been successfully led by Ts. Dr. Siti Azreena and her dedicated team, as well as the myHCI-UX Steering Committee members. We extend our heartfelt thanks for their time, effort, and leadership in organizing this year's event. All those hard work, vision, and commitment have ensured the continued excellence of FUSION.

Thus, we invite each and one of you to engage actively by sharing your work, challenging assumptions, collaborating across disciplines, and contributing to the future of human-computer interaction in Malaysia and beyond. Thank you for being part of FUSION 2025. Together, we can chart new pathways for impactful research, practice and education in HCI/UX.

Dr. Noris Mohd Norowi

President myHCI-UX 2025-27

Fusion 2025 Organization

Chair

Ts. Dr. Siti Azreena Mubin, Asia Pacific University of Technology & Innovation

Secretary

Ts. Muhammad Nur Affendy Bin Nor'a, Asia Pacific University of Technology & Innovation

Treasurer

Aida Zamnah Binti Zainal Abidin, Asia Pacific University of Technology & Innovation

Publication Unit

Ts. Dr. Siti Nurliana Jamali, University of Malaya

Dr. Shaima Abdulfattah Gamal Mohammed, Taiz University

Dr. Samia Abdhoo, Universiti Kebangsaan Malaysia

Khairunnisa Binti Misp, Brickfields Asia College

Communication & Promotion Unit

Ts. Siti Normaziah Binti Ihsan, University Malaysia Pahang Al-Sultan Abdullah

Dr. Tuty Asmawaty Binti Abd Kadir, University Malaysia Pahang Al-Sultan Abdullah

Dr. Siti Suhaila Binti Abd Hamid, University Malaysia Pahang Al-Sultan Abdullah

Noraniza Binti Samat, University Malaysia Pahang Al-Sultan Abdullah

Muhamad Adhwa Bin Mohd Salemi, University Malaysia Pahang Al-Sultan Abdullah

Mohd Hanafi Bin Hashim, Kolej Komuniti Pekan, Pahang

Registration Unit

Rizawati Binti Rohizan, Asia Pacific University of Technology & Innovation

Daniel Mago Vistro, Asia Pacific University of Technology & Innovation

Intan Nadiyah Abdul Hakim, German - Malaysia Institute

T Tulasi Sathyabama A/P Appalasaamy, Asia Pacific University of Technology & Innovation

Event and Technical Unit

Jeffery Jeselee Sijore, Asia Pacific University of Technology & Innovation

Ts. Dr. Siti Rahayu Abdul Aziz, University Technology Mara

SDC Unit (Student Design Challenge)

Aida Zamnah Binti Zainal Abidin, Asia Pacific University of Technology & Innovation

Tham Hoong Ching, Sunway College

Ts. Dr. Bavani Ramayah, University of Nottingham Malaysia

Ts. Dr. Hazwani Mohd Mohadis, International Islamic University Malaysia

Assoc. Prof. Ts. Dr. Vinothini Kasinathan, Asia Pacific University of Technology & Innovation

Assoc. Prof. Dr. Ing. Maizatul Hayati Mohamad Yatim, Universiti Pendidikan Sultan Idri

Table of Content

	Page
Foreword from Fusion 2025 Chair	i
Foreword from myHCI-UX President	ii
FUSION 2025 Organization	iii
Proceeding Paper	
Preferences of Mobile Learning Application Features Among Female Middle - Aged Learner to Learn AI-Quran <i>Nazean Jomhari, Syahida Mohtar, Baharul Hisyam, Mumtaz Begum Peer Mustafa Mohd Yakub @ Zulkifli Mohd Yusoff, Shamsul Arrieya Ariffin</i>	1
A Systematic Literature Review: Virtual Reality and Machine Learning for Alzheimer's Disease Diagnosis <i>Christian Nathaniel Tjandra, Siti Azreena Mubin, Khalifa Djemal, Amir Ali Feiz</i>	9
The Design of Web-Based Bachelor Project Assessment System using Design Thinking <i>Siti Athirah Othman, Nurul Saidahtul Fatihah Shaharuddin and Azrina Kamaruddin</i>	19
AI-Enhanced Hybrid Metaverse Framework for Multiplayer Interactive and Gamified Virtual Campus Tours <i>Sagar Prema Rajan, Siti Azreena Mubin, Vinesh Thiruchelvam</i>	28
Towards Considerate AI: Building a Fair and Polite Chatbot for Customer Service <i>Looi Pei Yi, Ng Jin Jin, Li ZhiHan, Raja Jamilah Raja Yusof, Viktor Kaptelinin</i>	37
Doctoral Consortium - Extended Abstract	
Extending Interaction Design Dimensions for Augmented Reality Supporting Children with Dyslexia <i>Emy Hazlinda Mohammad Ridzwan, Husniza Husni</i>	44
Social Media Platforms as Learning Environments: A Pilot Test of Discord and TikTok Live <i>Nurul Asmaa Abd Razak, Noris Mohd Norowi</i>	50
Collaborative Emotion-Recognition Brain-Computer Interface for User-Centric UX Evaluation <i>Ting Ying Lai, Masitah Ghazali</i>	61
Multivariate Determinants of Playability in Massively Multiplayer Online Game Teams: The Case of China <i>Xiaoxue Gong, Lili Nurliyana Binti Abdullah, Azrul Hazri Bin Jantan, Noris Binti Mohd Norowi</i>	66
Designing Co-Located Mixed Reality Gamelan Collaboration Using Non-verbal Cues <i>Mohamad Faris Hakim, Noris Mohd Norowi</i>	70

Preferences of Mobile Learning Application Features Among Female Middle - Aged Learner to Learn AI-Quran

Nazean Jomhari¹, Syahida Mohtar², Baharul Hisyam³, Mumtaz Begum Peer Mustafa⁴
Mohd Yakub @ Zulkifli Mohd Yusoff⁵, Shamsul Arrieya Ariffin⁶

^{1,3,4}Department of Software Engineering, Universiti Malaya, Kuala Lumpur

²Department of Software Engineering, Universiti Teknikal Malaysia, Melaka

⁵Department of Department of AI-Quran and AI-Hadith, Universiti Malaya, Kuala Lumpur

⁶Department of Computer Science and Digital Technology, UPSI, Perak

ABSTRACT

In the smartphone era, e-learning integration with mobile devices enables learning anytime and anywhere, known as mobile learning. While many studies explore mobile learning among college students, little is known about the readiness of middle-aged Muslims to adopt it for Quranic studies. This study investigates factors influencing the readiness of 168 Malaysian women aged 40 to 60, non-native Arabic speakers, to use mobile learning for Quranic Arabic. Data was collected via online questionnaires. Findings show that learners prefer specific features, particularly in-app communication and quizzes, which fulfill their needs for interaction and assessment-key to satisfaction and sustained engagement. Other valued features include tutor selection, push notifications, and offline access, though preferences differ by age, education, and Arabic knowledge. Less interest was shown in progress bars and gamification. The study concludes that effective mobile learning apps should emphasize interactivity and meaningful assessment tools.

Keywords: Readiness Factor, Mobile Learning, Middle-Aged, AI-Quran Learning

INTRODUCTION

Mobile learning readiness is an aspect of technology integration, and it is defined as the level of acceptance by teachers or students to adopt mobile learning. In the previous studies, readiness of teachers to adopt mobile learning in a classroom can be assessed using Mobile Learning Readiness Survey (MLRS) based on four-factor solution which are possibilities, benefits, preferences, and external influences (Christensen & Knezek, 2017), and mobile learning adoption by university students can be based on four fundamental constructs, which are technological, organizational, cultural, and quality factors (Almaiah & Al Mulhem, 2019). In the other hand, driving factors for adult learners in adopting mobile learning includes cognitive need, confirmation, satisfaction, subjective norms, perceived usefulness, perceived ease of use, and attitude (Hossain et al., 2021).

In Malaysia, there has been a steady growth in adult learners' involvement of Quran and Islamic learning activities. Traditionally, non-formal religious learning such learning activities were conducted in mosques, pondok schools, or madrasah by tradition. Adult learners who are interested in Quran and Islamic learning activities in Malaysia bears the following characteristics: spiritually driven, passion for knowledge, voluntary-learners,

associate experiences in learning (Bin Kadir et al., 2016). However, working adults with heavy responsibilities are usually unable to devote themselves to attend Quranic and Islamic class due to their time and location constraints. Therefore, this study aims to identify the factors related to the readiness of middle-aged adults to use mobile learning in the field of Quranic Arabic learning. 168 participants involved in this study were non-native Arabic speakers (N-NAS) who had no experience with daily conversation in Arabic and were confident in owning and using smartphones. In Malaysia, a study to determine the readiness of working adults in using mobile learning in Malaysia towards 97 of working adults aged between 18 and 40 years old, who volunteered to answer the online questionnaire shows that the respondents have adequate of knowledge in using mobile phone with internet connection, have a positive perspective toward integration of m-learning. However, they are not ready to utilize mobile learning (Zhao et al., 2021). This could be because the study was conducted before the Covid-19 pandemic where at that time many people had the opportunity to move freely to attend physical classes or perhaps the unreadiness may be due to the time and place constraints.

METHODOLOGY

This study is conducted in the scope of middle-aged Malay women who are not fluent in Arabic language as a first language, and who are comfortable enough to own or use mobile smartphones. The participants in this study consisted of 168 middle-aged women between the age of 40 to 60 years old. The participants were non-native Arabic Speakers (N-NAS), and at the same time did not have any experience in practicing daily conversational Arabic. Demographics and individual characteristics of the sample (age group, education level, occupation, and knowledge on learning Arabic language) are shown in Table 1.

Table 1: Demographics and individual characteristics of the samples

Characteristic Item		Number/Percentage
Gender	Female	168 (100%)
Age Group	40-45 Years Old	96 (57.1%)
	46-50 Years Old	36 (21.4%)
	51-55 Years Old	25 (14.9%)
	56-60 Years Old	11 (6.6%)
Education Level	Bachelor	86 (51.2%)
	Master	30 (17.9%)
	Diploma	26 (15.5%)
	STPM and below	16 (9.5%)
	PhD	10 (6%)
Occupation	Government Sector	72 (42.9%)
	Private Sector	60 (35.7%)
	Housewife	25 (14.9%)
	Self Employed	9 (5.4%)
	Not Disclosed	2 (1.2%)
Knowledge on Learning Arabic Language	Basic	83 (49.4%)
	None	78 (46.4%)
	Intermediate	7 (4.2%)

Responses from the participants were collected using online structured questionnaires via Google-form. The questionnaires are comprised of 55 items separated into eight sections, namely Section A-F. However, this paper focuses exclusively on Item E6 from Section E, which asked respondents: *“If Islamic Study learning materials were designed in the Mobile Portal, which features would you prefer to be included in the design of mobile learning application for Quranic Arabic learning?”*. This item aimed to identify users’ preferred

features for the design of a Quranic Arabic mobile learning application, providing valuable insights for future development and interface design considerations. Google form link containing the questionnaires were sent to randomly selected middle-aged female individuals using email, WhatsApp and Telegram App. The participants were informed that their responses will be anonymous, and the collected data will only be applied for research purposes. The participants were also notified that there are no right or wrong answers to the questionnaires. This perhaps will encourage the participants to provide sincere responses based on their thoughts, observations, and experiences.

RESULTS

Descriptive analysis to identify the preferences for mobile learning application features among the participants were performed using R. Items from the questionnaire enquired the participants' preferences for the features to be included in the design of mobile learning application for Quranic Arabic learning. The participants were provided with multiple-choice features consisting of choose tutor, In-App communication, reminder / push notification, offline learning, intuitive interface, quiz, progress bar, and gamification. Responses from the participants were then cross-referenced with their age group, education level, and basic Arabic knowledge for analysis. Figure 1 and Table 2 show the preferences for mobile application features based on participants age group. In-App communication was the favorite feature across all age groups with the highest percentage coming from 40-45 years old (n=83, 20.6%), followed by 51-55 years old (n=20, 18.7%), 46-50 years old (n=28, 17.7%), and 56-60 years old (n=8, 16.7%). Apart from In-App communication, the following were the outstanding top 3 features for each age group: 40-45 years old (quiz, gamification, choose tutor), 46-50 years old (push notification, quiz, choose tutor), 51-55 years old (quiz, choose tutor, push notification), and 56-60 years old (offline learning, quiz, push notification). For the least favorite feature, data shows that progress bar was the least favorite among the age group of 40-45 years old (n=31, 7.7%) and 46-50 years old (n=14, 8.9%), meanwhile intuitive interface was the least favorite among the age group of 51-55 years old (n=9, 8.4%) and 56-60 years old (n=3, 6.3%).

Breakdown for each feature revealed the following: In-App communication preferred by all age group; push notification preferred by 56-60 and 46-50 years old; offline learning preferred by 56-60 years old; intuitive interface preferred by 40-45 and 46-50 years old; quiz preferred by all age group; progress bar preferred by 56-60 years old; gamification preferred by 40-45 and 46-50 years old; and choose tutor preferred by 51-55 years old.

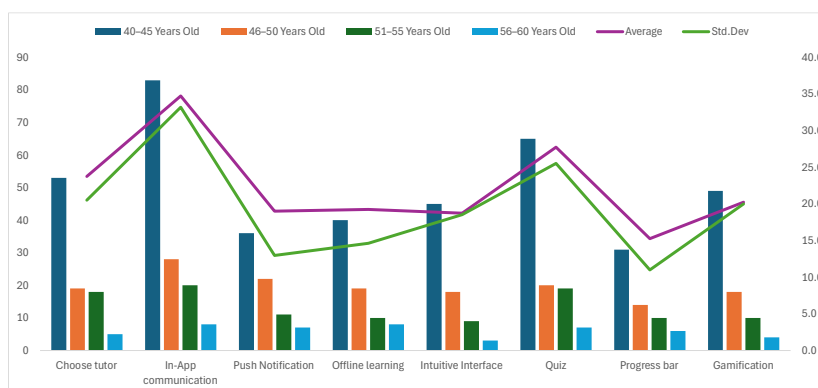


Figure 1: Preferences for mobile application features based on age group.

Table 2: Preferences for mobile application features based on age group.

Features	Age Group				Average	Std.Dev
	40-45 Years Old	46-50 Years Old	51-55 Years Old	56-60 Years Old		
Choose tutor	53 (13.2%)	19 (12%)	18 (16.8%)	5 (10.4%)	23.8	20.5
In-App communication	83 (20.6%)	28 (17.7%)	20 (18.7%)	8 (16.7%)	34.8	33.2
Push Notification	36 (9%)	22 (13.9%)	11 (10.3%)	7 (14.6%)	19.0	13.0
Offline learning	40 (10%)	19 (12%)	10 (9.3%)	8 (16.7%)	19.3	14.6
Intuitive Interface	45 (11.2%)	18 (11.4%)	9 (8.4%)	3 (6.3%)	18.8	18.6
Quiz	65 (16.2%)	20 (12.7%)	19 (17.8%)	7 (14.6%)	27.8	25.5
Progress bar	31 (7.7%)	14 (8.9%)	10 (9.3%)	6 (12.5%)	15.3	11.0
Gamification	49 (12.2%)	18 (11.4%)	10 (9.3%)	4 (8.3%)	20.3	20.0

Figure 2 and Table 3 show the preferences for mobile application features based on education level of the participants. In-App communication was the favorite feature across all groups except PhD holder and STPM and below holder, with the highest percentage came from diploma holder (n=24, 24.2%), followed by bachelor's holder (n=72, 21.2%), and master's holder (n=24, 17.1%). PhD holder favorite feature was quiz (n=10, 14.9%), and STPM and below holder favorite feature was offline learning (n=12, 17.1%). The following were the outstanding preferred features for each group: bachelor's holder (quiz, choose tutor, gamification), diploma holder (quiz, push notification), master's holder (choose tutor, quiz, gamification), PhD holder (intuitive interface, In-App communication), and STPM and below holder (choose tutor, In-App communication, push notification, quiz). For the least favorite feature, progress bar was the least favorite among bachelor's holder (n=28, 8.3%), diploma holder (n=6, 6.1%), and master's holder (n=13, 9.3%); meanwhile gamification was the least favorite among PhD holder (n=7, 10.4%) and STPM and below holder (n=5, 7.1%). Breakdown for each feature revealed the following: In-App communication preferred by all group; push notification preferred by diploma holder and STPM and below holder; offline learning preferred by STPM and below holder; intuitive interface preferred by PhD holder; quiz preferred by all group; progress bar preferred by

PhD holder; gamification preferred by master’s and degree’s holder; and choose tutor preferred by bachelor’s, master’s, and STPM and below holder.

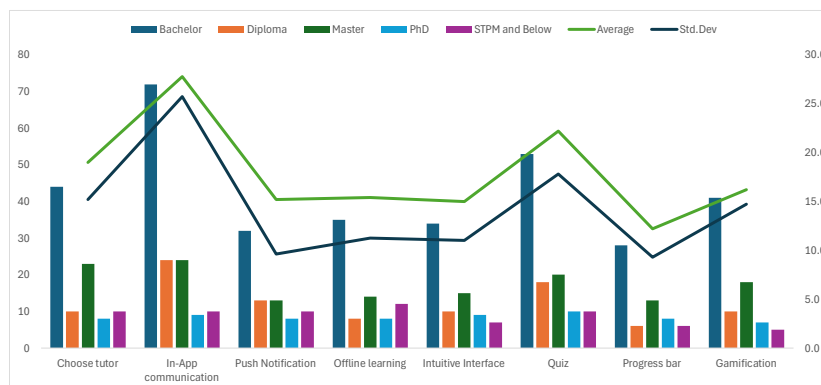


Figure 2: Preferences for mobile application features based on education level.

Table 3: Preferences for mobile application features based on education level.

Features	Education Level					Average	Std.Dev
	Bachelor	Diploma	Master	PhD	STPM and below		
Choose tutor	44 (13%)	10 (10.1%)	23 (16.4%)	8 (11.9%)	10 (14.3%)	19.00	15.20
In-App communication	72 (21.2%)	24 (24.2%)	24 (17.1%)	9 (13.4%)	10 (14.3%)	27.80	25.75
Push Notification	32 (9.4%)	13 (13.1%)	13 (9.3%)	8 (11.9%)	10 (14.3%)	15.20	9.63
Offline learning	35 (10.3%)	8 (8.1%)	14 (10%)	8 (11.9%)	12 (17.1%)	15.40	11.26
Intuitive Interface	34 (10%)	10 (10.1%)	15 (10.7%)	9 (13.4%)	7 (10%)	15.00	11.02
Quiz	53 (15.6%)	18 (18.2%)	20 (14.3%)	10 (14.9%)	10 (14.3%)	22.20	17.81
Progress bar	28 (8.3%)	6 (6.1%)	13 (9.3%)	8 (11.9%)	6 (8.6%)	12.20	9.28
Gamification	41 (12.1%)	10 (10.1%)	18 (12.9%)	7 (10.4%)	5 (7.1%)	16.20	14.72

Figure 3 and Table 4 show the preferences for mobile application features based on participants knowledge on learning Arabic language. The participants were classified into three groups, namely No, Basic, and Intermediate knowledge on learning Arabic language. From the analysis, In-App communication was the favorite feature across all groups except for Intermediate, with the highest percentage came from Basic group (n=74, 20.7%), followed by No group (n=60, 18.6%). Intermediate group preferred quiz as the favorite mobile application feature (n=7, 19.4%). The following were the outstanding top 3 features for each group: No (quiz, choose tutor, online learning), Basic (quiz, choose tutor, gamification), and Intermediate (gamification, In-App communication, progress bar). For the least favorite feature, progress bar was the least favorite among the No group (n=24, 7.5%) and Basic group (n=32, 9.0%), meanwhile choosing tutor was the least favorite feature for Intermediate group (n=1, 2.8%). Breakdown for each feature revealed the following: In-App communication preferred by all group; push notification, offline learning, and intuitive interface exhibited same level of interest among all group; quiz preferred by

all group; progress bar and gamification preferred by Intermediate group; and choose tutor preferred by No group.

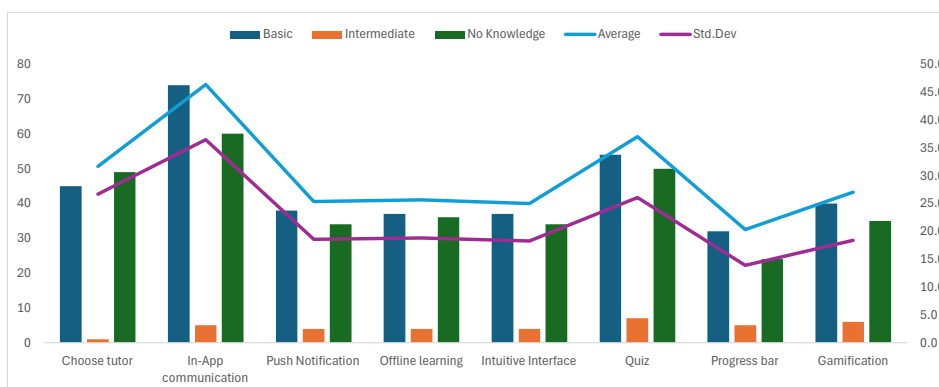


Figure 3: Preferences for mobile application features based on learning Arabic knowledge

Table 4: Preferences for mobile application features based on learning Arabic knowledge.

Features	Learning Arabic Knowledge			Average	Std.Dev
	Basic	Intermediate	No		
Choose tutor	45 (12.6%)	1 (2.8%)	49 (15.2%)	31.7	26.6
In-App communication	74 (20.7%)	5 (13.9%)	60 (18.6%)	46.3	36.5
Push Notification	38 (10.6%)	4 (11.1%)	34 (10.6%)	25.3	18.6
Offline learning	37 (10.4%)	4 (11.1%)	36 (11.2%)	25.7	18.8
Intuitive Interface	37 (10.4%)	4 (11.1%)	34 (10.6%)	25.0	18.2
Quiz	54 (15.1%)	7 (19.4%)	50 (15.5%)	37.0	26.1
Progress bar	32 (9%)	5 (13.9%)	24 (7.5%)	20.3	13.9
Gamification	40 (11.2%)	6 (16.7%)	35 (10.9%)	27.0	18.4

DISCUSSION AND CONCLUSION

This research aims to identify the preferences of mobile learning features among N-NAS middle-aged learners in adopting mobile learning. Eight mobile application features were presented to the respondents for selection as their preferred applications for inclusion in mobile learning. Analysis of the result demonstrated that across all groups (education level, age, and basic Arabic knowledge) in general, In-App communication and quiz were the most preferred mobile learning application features among middle-aged learners. This finding is consistent with the study conducted in Ilgaz and Gulbahar (2020) where learners' satisfaction in e-learning have a significant and positive relationship with "Interaction & Evaluation". In this case, providing In-App communication (Interaction property) and quiz (Evaluation property) features in mobile learning applications will considerably predict learners' satisfaction in using mobile learning.

In summary, Table 5 below shows the overall ranking of mobile learning application

features among middle-aged learners in adopting mobile learning to understand the Quran, and the average feature ranking based on respondents' age group, education level, and basic Arabic knowledge. The average ranking was obtained by comparing the preference scores of each feature within the three categories. Rank 1 indicates the most preferred feature. The overall ranking was then determined by averaging these three rankings and arranging the features in ascending order, with the lowest average value representing the highest overall preference.

Table 5: Ranking of feature preferences for mobile learning application among middle-aged learners

Features	Average Ranking			Overall Ranking
	Age Group	Education Level	Arabic Knowledge	
Choose tutor	3	3	4	3
In-App communication	1	1	1	1
Push Notification	3	4	6	5
Offline learning	3	4	5	4
Intuitive Interface	8	4	7	7
Quiz	2	2	1	2
Progress bar	7	8	8	8
Gamification	6	7	3	6

In the online learning environment where learning responsibility is managed by the learners, communication self-efficacy is an important element for overcoming the inadequacies of online communication (Hung et al., 2010). In-App communication could function as the tool of interaction and communication between learners and instructors. This result also supported by García-Martínez et al. (2019) where the study confirmed that the most widely used application in mobile learning focused on improving the interaction inside and outside the classroom. Quiz or projects/assignments whereas able to function as the support assessment for the course content. Study in Tawafak et al. (2018) reveals that course content significantly affects the support assessment, where it significantly affects the continued intention to use e-learning model. Performance assessment procedure should be maintained for adult learners to satisfy their cognitive need and continuous intention to use mobile learning (Hossain et al., 2021). Aside, the result from the analysis disclosed that the least preferred mobile application features across all groups in general was Progress bar. This finding is coherent with a study in Baldauf et al. (2017), where personal progress bar was not a popular feature in a mobile application for language learning.

In addition, it is important to consider the sociocultural context of middle-aged female learners who often balance learning with household and family responsibilities. For this group, features such as in-app communication and quizzes not only enhance engagement but also help reduce feelings of isolation in self-paced learning, providing a sense of connectedness and support. Furthermore, the relatively low preference for gamification and progress indicators suggests that overly playful or competitive elements may not align well with the seriousness and spiritual nature of Quranic learning. Therefore, developers should prioritize culturally respectful design choices that emphasize trust, ease of use, and meaningful interaction over entertainment-oriented features.

In the next phase, the level of association between mobile learning readiness factors and mobile learning aspects will be examined through analysis of variance, correlation, and linear regression.

REFERENCES

- Almaiah, M. A., & Al Mulhem, A. (2019). Analysis of the essential factors affecting of intention to use of mobile learning applications: A comparison between universities adopters and non-adopters. *Education and Information Technologies*, 24(2), 1433-1468.
- Baldauf, M., Brandner, A., & Wimmer, C. (2017, November). Mobile and gamified blended learning for language teaching: Studying requirements and acceptance by students, parents and teachers in the wild. In Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia (pp. 13-24).
- Bin Kadir, M. A., Arifin, S., & Fuad, A. N. (2016). Adult Learners' Understanding in Learning Islam Using Andragogy Approach: A Study in Kampung Siglap Mosque and Al-Zuhri Higher Learning Institute. *Journal of Education and Practice*, 7(32), 1-10.
- Christensen, R., & Knezek, G. (2017). Readiness for integrating mobile learning in the classroom: Challenges, preferences and possibilities. *Computers in human Behavior*, 76, 112-121.
- García-Martínez, I., Fernández-Batanero, J. M., Cobos Sanchiz, D., & Luque de La Rosa, A. (2019). Using mobile devices for improving learning outcomes and teachers' professionalization. *Sustainability*, 11(24), 6917.
- Hossain, M. N., Talukder, M. S., Khayer, A., & Bao, Y. (2021). Investigating the factors driving adult learners' continuous intention to use M-learning application: a fuzzy-set analysis. *Journal of Research in Innovative Teaching & Learning*, 14(2), 245-270.
- Hung, M.-L., Chou, C., Chen, C.-H., & Own, Z.-Y. (2010). Learner readiness for online learning: Scale development and student perceptions. *Computers & Education*, 55(3), 1080-1090.
- Ilgaz, H., & Gulbahar, Y. (2020). Examining e-learners' preferences and readiness satisfaction: A holistic modelling approach. *Open Praxis*, 12(2), 209-222.
- Tawafak, R. M., Romli, A. B., & Arshah, R. B. A. (2018). Continued intention to use UCOM: Four factors for integrating with a technology acceptance model to moderate the satisfaction of learning. *Ieee Access*, 6, 66481-66498.
- Zhao, S., Kinshuk, Yao, Y., & Ya, N. (2021). Adoption of mobile social media for learning among Chinese older adults in senior citizen colleges. *Educational technology research and development*, 69(6), 3413-3435.

A Systematic Literature Review: Virtual Reality and Machine Learning for Alzheimer's Disease Diagnosis

Christian Nathaniel Tjandra, Siti Azreena Mubin
Asia Pacific University of Technology and Innovation, Malaysia
Khalifa Djemal, Amir Ali Feiz
Paris-Saclay University, France

ABSTRACT

Alzheimer's Disease (AD) is a progressive neurodegenerative disorder where early detection is critical for timely intervention. However, traditional paper-based cognitive assessments lack ecological validity and rely on subjective clinician evaluation. This study conducts a systematic literature review (SLR) to examine how virtual reality (VR) cognitive assessments integrated with Machine Learning (ML) can enhance diagnostic accuracy and scalability. Following PRISMA 2020 guidelines, six empirical studies published between 2020 and 2025 were analysed for system design, biomarkers, ML models, and performance metrics. Results show that VR provides realistic, engaging cognitive tasks. At the same time, ML effectively processes multimodal datasets, including eye-tracking, hand-tracking, and behavioural data, achieving diagnostic accuracy above 89% and enabling predictions of disease progression. The review identifies key research gaps, including generalizability issues, limited cognitive domain coverage, a lack of clinical trials, VR accessibility challenges, and sensor noise. These findings highlight the potential of VR-ML systems as robust, objective tools for AD screening and severity classification.

Keywords: Alzheimer's Disease, Mild Cognitive Impairment, Virtual Reality, Machine Learning, Cognitive Assessment, Multimodal Biomarkers, Cognitive Parameters, Systematic Literature Review.

INTRODUCTION

Alzheimer's Disease (AD) is a progressive neurodegenerative disorder that impairs memory, thinking, and behaviour, with no cure available (Błaszczyk, 2020; Zvěřová, 2019). Early diagnosis is therefore critical for intervention as early as possible. Patients can be diagnosed through cognitive assessments where their mental abilities are measured to show any impairment due to AD.

Conventional cognitive assessments, though fast, cost-effective and reliable, have some limitations. First, cognitive assessments are practical mainly for detecting Mild Cognitive Impairment (MCI), which is the precursor to AD, but less reliable for diagnosing later stages (Alvarez-Sanchez et al., 2025). Second, results depend on the clinician's interpretation, which may result in subjective results and patients' stress during evaluation, potentially affecting the accuracy of their cognitive performance (Florian et al., 2020; Wong & Jacova, 2018). Third, paper-based tests emphasise memory and learning but ignore perceptual-motor and real-world tasks, thus lacking ecological validity (Padovani et al., 2025).

Virtual Reality (VR) offers a solution by creating immersive, naturalistic environments that enhance ecological validity and reduce patient stress (LoTempio et al., 2024). Studies show VR-based tasks can accurately reflect how cognitive functioning is utilised in daily activities (Cavedoni et al., 2020). Machine Learning (ML) further strengthens assessment by analyzing large multimodal datasets objectively, eliminating clinician bias, and improving accuracy (Aslan & Özüpak, 2025; Spooner et al., 2020). ML can process complex behavioural, visual and psychological data beyond human capacity, enabling earlier and more precise diagnosis (G. Tsang et al., 2020).

This research on Systematic Literature Review (SLR) therefore investigates existing studies on VR-ML approaches for AD diagnosis, with three objectives: (1) to comprehensively identify and synthesise existing research regarding VR-ML systems for AD diagnosis, (2) to evaluate the quality and reliability of evidence of existing VR-ML approaches on AD diagnosis, and (3) to identify research gaps and improvements for future research.

METHODOLOGY

For the SLR, the methodology of this section follows the PRISMA 2020. This SLR aims to provide a comprehensive understanding of how VR and ML technologies have been integrated as a multimodal approach for diagnosing AD, establishing a foundation for further research.

Inclusion and Exclusion Criteria

Inclusion criteria in the literature are:

1. Studies with multimodal VR+ML for AD diagnosis.
2. Journal Articles as the paper.
3. 5-year span (2020-2025).
4. Empirical data results of the system.

Exclusion criteria in the literature review are:

1. Non-Journal Articles.
2. Non-English paper.
3. Not Accessible.
4. Research that utilises a single solution, either VR or ML, for AD diagnosis (supporting literature).

STUDY SELECTION PROGRESS

To obtain the relevant studies for data extraction and analysis, literature searches were conducted across several databases, including PubMed, IEEE Xplore, and Google Scholar, in addition to manual searching. Search terms that were used are ("Alzheimer's Disease" OR "Alzheimer Disease" OR "Dementia" OR "Mild Cognitive Impairment" OR "MCI") AND ("severity" OR "disease progression" OR "staging" OR "grading" OR "classification" OR "conversion" OR "early detection") AND ("virtual reality" OR "VR" OR "immersive" OR "immersive technology") AND ("machine learning" OR "deep learning" OR "artificial intelligence" OR "AI" OR "neural network*" OR "support vector machine" OR "transfer

learning"). This search strategy is structured to ensure that only relevant studies will appear in the search for the screening, which is done by 20 June 2025.

The study selection process will be conducted in several stages. First, the search results from the database will be selected based on the key terms used in the search. Second, the filters will be applied to limit the results to only journal articles published between 2020 and 2025. Third, the screening process will involve reviewing the research title and abstract to identify studies relevant to the SLR. The final step involves a full-text review, which entails a thorough assessment of the literature content to confirm its inclusion in the SLR.

RESULTS

A total of 4316 records were screened for the SLR (Figure 1). Following the entire process of screening, 51 journals were selected to go through the whole review process. After the full review process, a total of 6 studies were included, with data extracted. Forty additional studies will support the literature and enrich the content of the SLR after review.

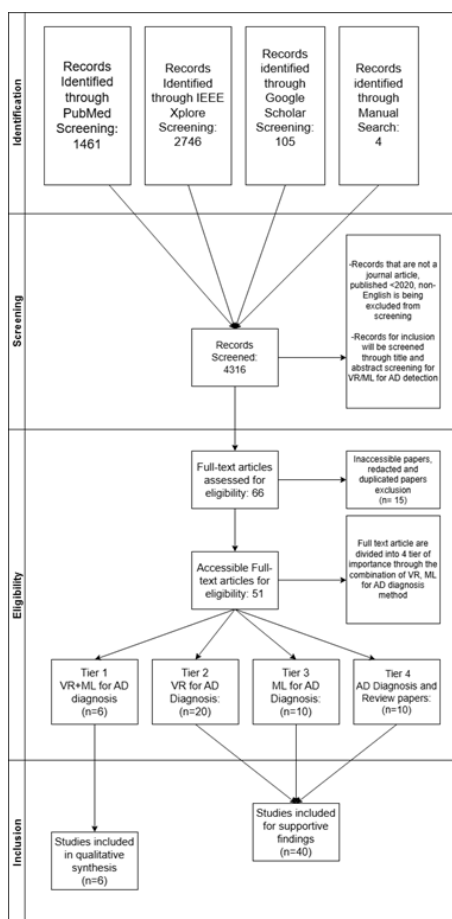


Figure 1. SLR Inclusion and Exclusion Process Diagram

TABLE OF DATA EXTRACTION

To systematically compare the reviewed studies, key information was extracted, including study type, sample size, VR system, additional biomarkers, cognitive parameters, ML

models, and main findings. Table 1 summarizes these characteristics, providing a clear overview of how different VR-ML approaches were applied to AD diagnosis.

Table 1. SLR data extraction

References	Study Type	Target Population	VR System	Extra Biomarker	Cognitive Parameters	ML Models	Conclusion
(Kim et al., 2023)	First Validation Study	51 patients (31 MCI and 20 healthy)	Virtual Kiosk Task	Hand tracking, eye-tracking	Language, attention, visuospatial, memory, executive function	Support Vector Machine (SVM)	Accurate in separating healthy and MCI patients
(Park et al., 2024)	Validation Study	54 patients (32 MCI and 22 healthy)	Virtual Kiosk Task	Hand tracking, eye-tracking	Language, attention, visuospatial, memory, executive function	Support Vector Machine (SVM)	Accurate diagnosis with more diverse biomarkers for better accuracy.
(Tsai et al., 2021)	Quasi-Experimental and Case-Control design	12 Patients (6 MCI and 6 healthy)	Virtual Supermarket	Behavioral movement and time history trajectory	Attention, language, visuospatial, memory, executive function	Logistic Regression, Support Vector Machine (SVM), Decision Tree, Random Forest, AdaBoost, and XG Boost.	Nearly perfect results in diagnosing and distinguishing between healthy and MCI.
(Wang et al., 2024)	Observational study	1328 patients (301 MCI and 1027 healthy)	Virtual Supermarket	Behavioral movement and time trajectory history	Memory, visuospatial, executive function, computational, attention	KM-Borderline SMOTE-ACWUEL	Demonstrated that integrating VR with ML enables ecological validity and scalable diagnosis.
(Wu et al., 2023)	Quasi-experimental study	86 patients (42 healthy and 44 MCI)	VR Language test	EEG, exploration pattern, time and speech latency	Language	SVM-BRF	Better user engagement and accurate tools for



				and fluency			early detection
(Xu et al., 2024)	Validation Study	201 patients with MCI	VECA (VR Eye tracking Cognitive Assessment)	Eye tracking	Memory, attention, calculation, abstraction, executive, language, and visuospatial	SVR (Support Vector Regression)	VECA is a reliable and non-invasive screening tool for AD

VR AS A TOOL FOR COGNITIVE ASSESSMENT

All studies highlight VR as a powerful tool for cognitive assessment, offering greater engagement and ecological validity than traditional paper-based tests (Wang et al., 2024; Wu et al., 2023). Using head-mounted displays, patients complete immersive tasks that mirror real-life situations, providing a clearer view of how cognitive abilities translate into daily functioning.

Examples include the Virtual Kiosk Test, which detects memory lapses and executive errors during ordering and payment tasks (Kim et al., 2023; Park et al., 2024), and supermarket simulations that assess planning, attention, and decision-making while navigating a shopping list (Tsai et al., 2021; Wang et al., 2024). Other tasks focus on language through verbal descriptions in virtual environments (Wu et al., 2023) or broader domains in the VECA system, which tests attention, memory, and problem-solving (Xu et al., 2024).

By aligning with established tools like MoCA, VR assessments remain clinically valid while replicating real-world demands. This ability to blend structured testing with immersive, interactive experiences demonstrates VR's potential to transform conventional cognitive assessments.

ML Integration for AD Diagnosis

VR simulations combined with biomarkers generate large datasets that require ML for analysis. Unlike traditional statistical methods, ML captures non-linear relationships and heterogeneous inputs, with ensemble models adapting to varied data for higher diagnostic accuracy (Spooner et al., 2020; Tsai et al., 2021; Wang et al., 2024). Beyond classification, ML can also predict disease progression. Hybrid approaches integrating unsupervised and supervised methods improved predictive accuracy by 6.5% through mapping patient trajectories over time (Buchholz et al., 2023). Such predictive power supports early intervention and personalized care (Chudzik et al., 2024).

Most of the studies conduct comparative analysis across support vector machines (SVM), tree-based ensembles, and deep learning models, revealing that performance within the VR-ML system diagnosis depends on the dataset characteristics. Studies with SVMs consistently deliver high accuracy percentage (90-94%) on small or medium-sized datasets

because of the capability to handle limited but complex datasets. On the other hand, tree-based models like random forest and adaptive ensembles work better with large and uneven datasets, as they are able to balance the results of different groups and stabilize without intensive tuning. Deep learning architectures may be suitable for unstructured and complex datasets from EEG or VR motion data but may suffer from overfitting on small samples. Overall, SVM and ensemble methods remain as the most reliable for VR-ML systems.

ML performance is typically evaluated using accuracy, AUC and F1-score. Accuracy measures correct classifications, AUC reflects discrimination power, and F1 balances precision and recall for imbalanced datasets (Tsai et al., 2021; Wang et al., 2024). Together, these metrics provide a robust assessment of ML effectiveness in distinguishing cognitive impairment. The data of the studies with accuracy, AUC and F1 scores can be referred to in Table 2.

Table 2. VR-ML diagnosis results

References	Sample Size	Data Source	Accuracy	AUC Scores	F1 Scores
(Kim et al., 2023)	51	Task scores + completion time + eye-tracking + hand-tracking	93.90%	0.98	0.947
(Park et al., 2024)	54	Task scores + completion time + eye-tracking + hand-tracking + MRI	94.40%	0.93	0.957
(Tsai et al., 2021)	12	Navigation trajectory + movement pattern + task scores	100%	1	1
(Wang et al., 2024)	1328	Task scores + distance moved + completion time	92.46%	0.928	0.928
(Wu et al., 2023)	86	EEG + speech data + time duration	89.9%	-	0.87
(Xu et al., 2024)	202	Task-scores + eye-tracking data + demographic data	90%	0.95	-

MULTIMODAL DIGITAL BIOMARKERS

All the reviewed studies incorporated multimodal biomarkers to complement VR cognitive assessments to improve diagnostic precision. These biomarkers provide richer datasets that go beyond conventional task scores, enabling the detection of subtle behavioural and psychological changes linked to cognitive impairment. By integrating these inputs into ML frameworks, researchers were able to analyse high-volume, complex datasets and generate more accurate evaluations of cognitive function. This multimodal approach ensures that impairments are not only assessed through test outcomes but also through the way patients interact with and respond to their environment.

Among the most widely used measures, eye tracking and hand tracking offered detailed insights into attention, visuospatial ability, and motor control. Eye tracking, applied in three studies, revealed gaze behaviour, fixation duration, and visual observation strategies, consistently showing that MCI patients were more easily distracted and less focused than

healthy controls (Kim et al., 2023; Park et al., 2024; Wu et al., 2023). Hand tracking, included in two studies, captures movement trajectories during tasks such as object selection or kiosk interactions. These data highlighted how irregular or inefficient motion patterns could signal underlying executive and visuospatial deficits, providing a layer of diagnostic evidence unavailable in paper-based assessments (Park et al., 2024).

Other studies expanded on behavioural tracking by analysing navigation paths, task completion times, and exploration patterns within VR environments (Tsai et al., 2021; Wang et al., 2024; Wu et al., 2023). Unlike traditional assessments that only record correct and incorrect responses, this approach considered how patients approached tasks, the strategies they employed, and the time required to finish. Such data allowed researchers to detect these multimodal biomarkers with ML analysis. The reviewed studies demonstrated a more comprehensive framework for understanding cognitive decline, one that captures both outcome accuracy and the nuanced processes leading to those outcomes.

DISCUSSION

This SLR analysed six studies that integrated VR cognitive assessments with ML for diagnosing AD. The results consistently demonstrated high diagnostic performance, with accuracy rates around 90% (Kim et al., 2023; Park et al., 2024; Xu et al., 2024). These findings highlight the strength of combining immersive VR environments, which replicate real-world tasks and enhance ecological validity, with ML models capable of objectively processing complex multimodal datasets.

The reviewed studies show that VR tasks such as supermarket navigation or kiosk interactions provide richer insights into cognitive decline than traditional paper-based tests. Unlike written or verbal recall, VR-based activities simulate daily life challenges, revealing deficits in memory, attention and executive functioning. ML further improves diagnostic accuracy by analyzing diverse inputs, including eye and hand tracking, EEG and MRI, beyond human capabilities. Notably, some models demonstrated predictive ability, allowing not only the detection of MCI but also the forecasting of disease progression. This dual role strengthens the clinical relevance of VR-ML systems.

Despite promising outcomes, VR systems may not suit patients with severe impairment, sensory limitations, or motion sickness, raising accessibility concerns (Wang et al., 2024). Furthermore, most studies lack large-scale clinical trials and expert-driven validation, as many tools were built before neuropsychological consultation. Another limitation is the narrow focus on MCI, leaving later severity stages of AD stages unexplored. Finally, sensor reliability issues such as calibration drift and motion latency introduce noise that can affect ML accuracy.

Ethical and accessibility considerations are also crucial for the responsibility of VR-ML system testing on a clinical setting. All types of multimodal data must be protected through confidential and privacy collection process. Algorithmic bias is also important to be considered due to small number or unbalanced datasets which may produce less accuracy for the patients. Accessibility should be improved by VR systems that have is more user-friendly towards more severe cognitive levels, which could reduce discomfort and motion

sickness. This could improve the trust of the patients, by safeguarding privacy and inclusive design principles to ensure that the VR-ML diagnostic system is clinically safe and socially responsible.

The novelty of these studies lies in demonstrating how VR-ML systems can surpass existing cognitive assessments by combining realistic cognitive assessment and computational accuracy. Unlike traditional methods, these systems can simultaneously capture behavioural, physiological, and cognitive data, producing more comprehensive clinical profiles. This VR-ML system is not only a diagnostic tool but also a potential means to monitor disease progression over time.

FUTURE RESEARCH

This SLR shows that VR combined with ML offers a strong foundation for AD screening. Future work should extend severity detection beyond MCI to late AD stages and enrich datasets with additional biomarkers such as eye-tracking and hand tracking, enabling both classification and progression detection. Expert validation will be essential to improve generalizability, identify critical cognitive domains, and enhance system comfort and accessibility, ensuring tools are clinically relevant and user-friendly. Expert validation here is very important, due to the VR simulation system might have different measurements of the biomarkers compared to existing cognitive tests, to ensure its clinical validation. This also calls for the importance of large-scale clinical trials which could evaluate the accuracy of the system based on real patient's test. Finally, technical refinements are needed to reduce sensor noise through better calibration, filtering, and redundancy checks. These improvements will strengthen diagnostic reliability and support broader clinical adoption of VR-ML systems.

CONCLUSION

The SLR demonstrates that integrating VR cognitive assessments with ML provides a highly accurate and scalable approach for AD diagnosis. The originality of this research lies in highlighting VR's ability to create ecologically valid assessments and ML's strength in analyzing complex multimodal datasets to detect cognitive impairment and predict disease progression. The review confirms that VR-ML systems surpass traditional paper-based methods by enhancing engagement, reducing subjectivity, and enabling early detection. These findings establish a concrete foundation for advancing diagnostic tools toward classifying later disease stages and guiding timely interventions.

ACKNOWLEDGEMENT

The authors would like to thank Asia Pacific University of Technology and Innovations (APU) for providing a research grant (RDIG/02/2024) and MyTIGER 2024 (Embassy of France in Malaysia) for this study. The authors also acknowledge the anonymous reviewers for their valuable feedback and comments.

REFERENCES

- Alvarez-Sanchez, L., Pereto, M., Garcia-Valles, L., Balaguer, A., Pena-Bautista, C., Ferrer-Gonzalez, L., Baquero, M., & Pericas, C. C. (2025). Fast Declining Prediction in Alzheimer's Disease from Early Clinical Assessment. *Current Neuropharmacology*, 23(5), 602–611. <https://doi.org/10.2174/011570159X332930240925095423>
- Aslan, E., & Özüpak, Y. (2025). Comparison of machine learning algorithms for automatic prediction of Alzheimer disease. *Journal of the Chinese Medical Association : JCMA*, 88(2), 98–107. <https://doi.org/10.1097/JCMA.0000000000001188>
- Błaszczuk, J. W. (2020). Energy Metabolism Decline in the Aging Brain—Pathogenesis of Neurodegenerative Disorders. *Metabolites*, 10(11), 450. <https://doi.org/10.3390/metabo10110450>
- Bucholc, M., Titarenko, S., Ding, X., Canavan, C., & Chen, T. (2023). A hybrid machine learning approach for prediction of conversion from mild cognitive impairment to dementia. *Expert Systems with Applications*, 217, 119541. <https://doi.org/10.1016/j.eswa.2023.119541>
- Cavedoni, S., Chirico, A., Pedroli, E., Cipresso, P., & Riva, G. (2020). Digital Biomarkers for the Early Detection of Mild Cognitive Impairment: Artificial Intelligence Meets Virtual Reality. *Frontiers in Human Neuroscience*, 14, 245. <https://doi.org/10.3389/fnhum.2020.00245>
- Chudzik, A., Śledzianowski, A., & Przybyszewski, A. W. (2024). Machine Learning and Digital Biomarkers Can Detect Early Stages of Neurodegenerative Diseases. *Sensors*, 24(5), 1572. <https://doi.org/10.3390/s24051572>
- Florian, M., Margaux, S., & Khalifa, D. (2020). Cognitive tasks modelization and description in VR environment for Alzheimer's disease state identification. *2020 Tenth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 1–7. <https://doi.org/10.1109/ipta50016.2020.9286627>
- G. Tsang, X. Xie, & S. -M. Zhou. (2020). Harnessing the Power of Machine Learning in Dementia Informatics Research: Issues, Opportunities, and Challenges. *IEEE Reviews in Biomedical Engineering*, 13, 113–129. <https://doi.org/10.1109/RBME.2019.2904488>
- Kim, S. Y., Park, J., Choi, H., Loeser, M., Ryu, H., & Seo, K. (2023). Digital Marker for Early Screening of Mild Cognitive Impairment Through Hand and Eye Movement Analysis in Virtual Reality Using Machine Learning: First Validation Study. *Journal of Medical Internet Research*, 25, e48093. <https://doi.org/10.2196/48093>
- LoTempio, S., Johnson, S., Rice, M., Masters, R., Collins, S.-A., Hoefcker, J., Rivera, J., Schreiber, D., Interrante, V., Ortega, F., & Davalos, D. (2024). Aging Naturally: Virtual reality nature vs real-world nature's effects on executive functioning and stress recovery in older adults. *2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 1061–1062. <https://doi.org/10.1109/VRW62533.2024.00327>
- Padovani, A., Caratozzolo, S., Galli, A., Crosani, L., Zampini, S., Cosseddu, M., Turrone, R., Zancanaro, A., Gumina, B., Vicini-Chilovi, B., Benussi, A., Vyshedskiy, A., & Pilotto, A. (2025). Validation and convergent validity of the Boston cognitive assessment (BOCA) in an Italian population: A comparative study with the Montreal cognitive assessment (MoCA) in Alzheimer's disease spectrum. *Neurological Sciences : Official Journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*, 46(2), 697–704. <https://doi.org/10.1007/s10072-024-07775-3>
- Park, B., Kim, Y., Park, J., Choi, H., Kim, S.-E., Ryu, H., & Seo, K. (2024). Integrating Biomarkers From Virtual Reality and Magnetic Resonance Imaging for the Early Detection of Mild Cognitive Impairment Using a Multimodal Learning Approach: Validation Study. *Journal of Medical Internet Research*, 26, e54538. <https://doi.org/10.2196/54538>

- Spooner, A., Chen, E., Sowmya, A., Sachdev, P., Kochan, N. A., Trollor, J., & Brodaty, H. (2020). A comparison of machine learning methods for survival analysis of high-dimensional clinical data for dementia prediction. *Scientific Reports*, *10*(1), 20410. <https://doi.org/10.1038/s41598-020-77220-w>
- Tsai, C.-F., Chen, C.-C., Wu, E. H.-K., Chung, C.-R., Huang, C.-Y., Tsai, P.-Y., & Yeh, S.-C. (2021). A Machine-Learning-Based Assessment Method for Early-Stage Neurocognitive Impairment by an Immersive Virtual Supermarket. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *29*, 2124–2132. <https://doi.org/10.1109/TNSRE.2021.3118918>
- Wang, Y., Yang, P., Yu, J., Zhang, S., Gong, L., Liu, C., Zhou, W., & Peng, B. (2024). An Ensemble Learning Algorithm for Cognitive Evaluation by an Immersive Virtual Reality Supermarket. *IEEE Transactions on Neural Systems and Rehabilitation Engineering : A Publication of the IEEE Engineering in Medicine and Biology Society*, *32*, 3761–3772. <https://doi.org/10.1109/TNSRE.2024.3470802>
- Wong, S., & Jacova, C. (2018). Older Adults' Attitudes towards Cognitive Testing: Moving towards Person-Centeredness. *Dementia and Geriatric Cognitive Disorders Extra*, *8*(3), 348–359. <https://doi.org/10.1159/000493464>
- Wu, R., Li, A., Xue, C., Chai, J., Qiang, Y., Zhao, J., & Wang, L. (2023). Screening for Mild Cognitive Impairment with Speech Interaction Based on Virtual Reality and Wearable Devices. *Brain Sciences*, *13*(8). <https://doi.org/10.3390/brainsci13081222>
- Xu, Y., Zhang, C., Pan, B., Yuan, Q., & Zhang, X. (2024). A portable and efficient dementia screening tool using eye tracking machine learning and virtual reality. *NPJ Digital Medicine*, *7*(1), 219. <https://doi.org/10.1038/s41746-024-01206-5>
- Zvěřová, M. (2019). Clinical aspects of Alzheimer's disease. *Clinical Biochemistry*, *72*, 3–6. <https://doi.org/10.1016/j.clinbiochem.2019.04.015>

The Design of Web-Based Bachelor Project Assessment System using Design Thinking

Siti Athirah Othman, Nurul Saidatul Fatiha Shaharuddin and Azrina Kamaruddin
Universiti Putra Malaysia, Selangor

ABSTRACT

This paper presents the high-fidelity design development of FYPAAssess, a web-based Bachelor Project Assessment System developed using the Design Thinking approach. The study highlights the inefficiencies of the current manual process involving coordinators, supervisors, assessors, and students, which creates an administrative burden, time inefficiency and restricted access to mark compilation and calculation. FYPAAssess is designed to overcome these challenges by enabling seamless coordination including student assignment to supervisors, mark input by supervisors and assessors, automated calculation, and integrated notification features. The FYPAAssess design emphasis is placed on a user-friendly interface guided by principles of simplicity, clarity, and familiarity to enhance usability by implementing the Design Thinking approach. The proposed solution aims to streamline evaluation workflows, improve communication among stakeholders, and strengthen the overall management of the Bachelor Project assessment process.

Keywords: Web-based Bachelor Project, Evaluation, Assessment Automation, Design Thinking, Prototype

INTRODUCTION

This paper addresses the challenges of the manual process associated in managing and accessing the Bachelor Project, a compulsory two-semester course as outlined in the Undergraduate Study Programme Hand Book 2022-2023 Faculty of Computer Science and Information Technology (FCSIT). The course comprises Part A (SWE4949A) in Semester 6 and Part B (SWE4949B) in Semester 7, with multiple assessment types evaluated by supervisors and assessors. Besides, students are evaluated across a range of criteria, including communication skills, domain knowledge, project demonstration, digital competencies, proposal quality, and the commercial viability of their projects, all of which align with the course learning objectives. These criteria are structured according to Bloom's Taxonomy and tied closely with the assessment rubric (Alammary & Masoud,2025).

Thus, assessing Bachelor Project requires many aspects and criteria to be assessed, whereas in our context, the assessment methods are still relying on paper-based forms and Excel-dependent processes by the supervisor, assessor and coordinator. To address this, the objective of this study is to design a web-based system, FYPAAssess, that automates assessment management by providing a centralized platform tailored to the roles of coordinators, supervisors, assessors, and students, thereby streamlining workflows and improving transparency using Design Thinking (DT).

BACKGROUND

Based on Body of Knowledge (BOK) in Computer Science, by implementing project bachelor, students are hoping to present certain marketable competency when they can apply the knowledge, skills and disposition in context (CC2020 Task Force, 2020) from each course from their early semester to the end of semester in SE (Hironor, 2024) or Computer Science (CS). As a result, project bachelor project course carries six credit hours and is compulsory for all undergraduate students. It is divided into two parts, SWE4949A and SWE4949B, which are offered in semester 6 and semester 7 for students to take. Moreover, the assessment rubrics for this project bachelor evaluation must explicitly mapped with degree program learning outcomes.

At FCSIT, UPM the process or project bachelor evaluation involved coordinator, supervisor and assessor. The department project bachelor coordinator is involved in managing the whole project bachelor administrative work coordinated fluently. Many years, we used paper-based form to input student marks. Next, the coordinator gathered the paper-based forms to input each student marks from supervisor and assessors in Excel spreadsheet according to the course objectives stated. Due to these manual tasks by coordinator, supervisor and assessors collecting and giving marks to students, it will produce many inefficiencies in completing the task. The problems can come from all user personas: coordinator, supervisor and assessors such as miscalculation student marks, incorrect data input by all three user personas and backlog tasks where coordinator needs to publish the student marks to their supervisor. Nonetheless, we have identified that the challenges that we have also happen to other institutions (Kusay, 2023; Kar et al., 2017; Thompson & Ahn,2012).

Project Bachelor Assessment

Assessments are central to student learning and skill development. Effective assessments target multiple cognitive levels, measure diverse abilities, and align with course learning outcomes (CLOs). Instructors must therefore classify both assessment tasks and CLOs to ensure quality and enhance the learning experience (Alammary & Masoud,2025). Regardless of any bachelor programme from any institution the assessment CLOs and its management vary. In some institution the management of assessment for example: mark input by supervisor and assessor are using the system.

Existing systems to assess project bachelor

Several universities already have digital systems to help manage and improve the process of final year project assessments, and these give useful ideas for FYPAAssess . At Victoria University of Wellington, the ECS Final Year Project Marking System tries to solve problems like examiner assignment and grading transparency by giving role-based views and automatic progress tracking (Kusay, 2023).

UiTM's oPENs system brings project documents into one place and gives role-based access for administrators, coordinators, lecturers, students, and panels. It also allows FYP titles to be

edited and uses rubric-based evaluations, but it still depends on other systems like SIMS for submitting final marks (Kar et al., 2017).

KAIST's Online Grading System supports distributed grading with outlier removal and rubric scoring, which fits well with SWE4949's two-assessor structure (Thompson & Ahn, 2012). Programming Assessment Grading System by Edupuganti makes grading programming tasks easier by letting professors assign grades inside the system, but since it cannot be edited and has a narrow scope, it is not suitable for other academic areas (Edupuganti, 2013).

The eFYP Management System from Affandi et al. (2022) is mobile-based and supports document centralization and progress tracking, but it does not have a grading module, so marks still need to be entered manually into SPMP.

METHODOLOGY

To provide the best design solution, the Design Thinking (DT) approach was adopted to ensure that user requirements were accurately identified and addressed. The target users of the systems include coordinators, supervisors and assessors, primarily lecturers from the FCSIT, UPM, as well as students enrolling in Project Bachelor. We will follow the DT approach to solve this problem. The approaches are empathy, define, ideate, prototype and test.

Case study

At one of the faculty in University Putra Malaysia (UPM) which the faculty comprises four departments: computer system, software engineering & information system, network and communication and multimedia. The organization has been using the *Sistem Maklumat Pelajar* or e-SMP to input all course marks. However, the e-smp can only be accessed by the project coordinator to input the project bachelor marks. Thus, the mark given and calculated before the coordinator can input in the system were by paper-based forms and Excel spreadsheet.

Participant

At the FCSIT we have four departments. In each department one FYP coordinator is appointed to assist each department's needs and requirements. We will invite four coordinators, four supervisors, four assessors and eight students from various departments at FCSIT. The interview will be conducted depending on participants' options, it can be via online or physical.

The Analysis

Empathy map and user journey map were used to identify the user persona and to understand the user pain, gain and aims. Next, with the usage of these analysis tools, we will identify user persona and understand our user persona better before we proceed with the solution (Ning, 2025).

Materials

To acquire input and analysis about our users, we used many tools such as Canva and Figma to assist the analysis done within the development team members. The tools were selected because it supports collaborative tasks.

RESULT AND FINDINGS

User Persona

From the analysis we found four user's personas: coordinator, supervisor, coordinator and student.

Each of the user persona's shared similar defined problem about inefficiency in managing student's mark under their supervision and examining the student under their supervision and other student. The existing process flow has been used paper-based form by supervisors and assessors to input marks. Then the coordinator collected all the paper-forms marks and input all marks from supervisors and assessors in the Excel format template. As a result, there is a high potential human-error that can occur when managing marks using manual input from supervisors, assessor and coordinator. This empathy and define phases have allowed the researcher to understand the process flow about the project bachelor marks managing process and issues and challenges from four user's personas.

Ideation and Prototyping

When the researcher has a clear understanding about the defined problem of four user's personas, we analyzed the ideation process and decided to provide a web-based solution to the problems. Once we have a clear understanding of the problems and type of solution that we are going to build, a case diagram as shown in figure 1 was developed to visualize the functional scope of the proposed system.



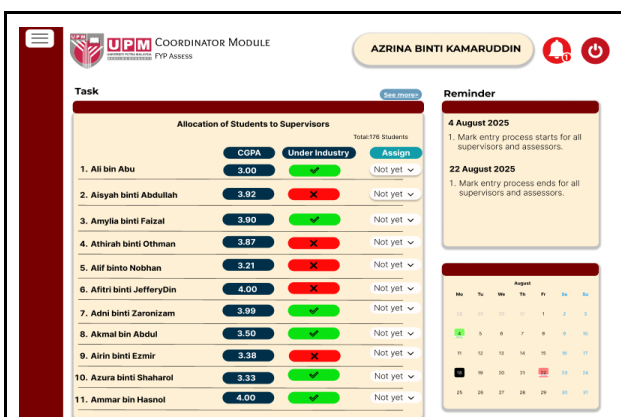
Figure 1. Use case diagram

Next phase is to produce low-fidelity prototype. All user interfaces (UI) mock-up's first version are produced for the web-based *FYPAssess* such as figures 2, 3, and 4 illustrate the dashboards for the coordinator, assessor, and supervisor modules, which provide users with quick access to information.

Next, first version of UI mock-ups for student user persona are: figure 5 is the CGPA and resume submission by the students, figure 6 is about the submission of project title by student. Figure 7, the mark entry UI offers direct access to simplified evaluation criteria derived from the assessment rubric to facilitate users' understanding in determining students' marks.

Figure 8 the UI demonstrates shows the form that the supervisor needs to state whether they have a collaboration with the industry or not. Figure 9 is the UI for supervisors to view the report about their students under their supervision. Figure 10 shows the student assignment page. This is for the coordinator to allocate the students to the supervisors, where the coordinator can choose either to allocate the students manually or automatically, which is based on the AI recommendation. The notification page, as in Figure 11, shows all the tasks that need to be completed before the deadline by the supervisors and assessors. This is to ensure the time management of the Bachelor Project's course is handled well without any time extension. Figure 12 allows the coordinator to monitor the mark submissions made by the assessors and supervisors. Under this section as well, coordinators are allowed to notify the supervisors and assessors to complete the tasks as they are approaching the deadline. Lastly, figure 13 allows

the coordinators to modify the learning objectives of the course. As the learning objectives are subject to change, this simplifies the modification of learning objectives.



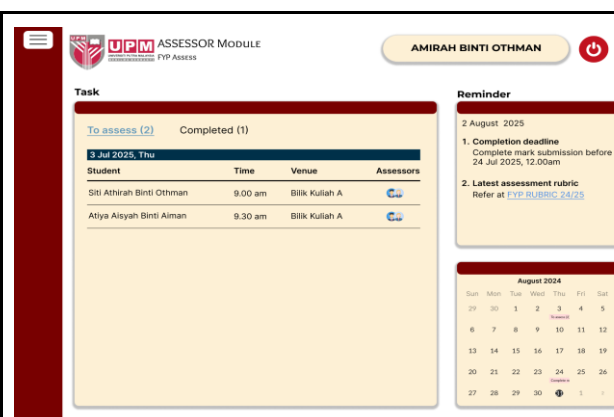
UPM COORDINATOR MODULE
AZRINA BINTI KAMARUDDIN

Task
Allocation of Students to Supervisors
Total: 176 Students

No.	Name	CGPA	Under Industry	Assign
1.	Alli bin Abu	3.00	Not yet	Assign
2.	Aisyah binti Abdullah	3.92	X	Not yet
3.	Amylia binti Faizal	3.00	Not yet	Not yet
4.	Athirah binti Othman	3.87	X	Not yet
5.	Alif binto Nobhan	3.21	X	Not yet
6.	Alfitri binti JefferyDin	4.00	X	Not yet
7.	Adni binti Zaronizam	3.99	Not yet	Not yet
8.	Akmal bin Abdul	3.50	Not yet	Not yet
9.	Airin binti Ezmir	3.38	X	Not yet
10.	Azura binti Shaharol	3.33	Not yet	Not yet
11.	Ammar bin Hasnol	4.00	Not yet	Not yet

Reminder
4 August 2025
1. Mark entry process starts for all supervisors and assessors.
22 August 2025
1. Mark entry process ends for all supervisors and assessors.

Figure 2. Coordinator dashboard



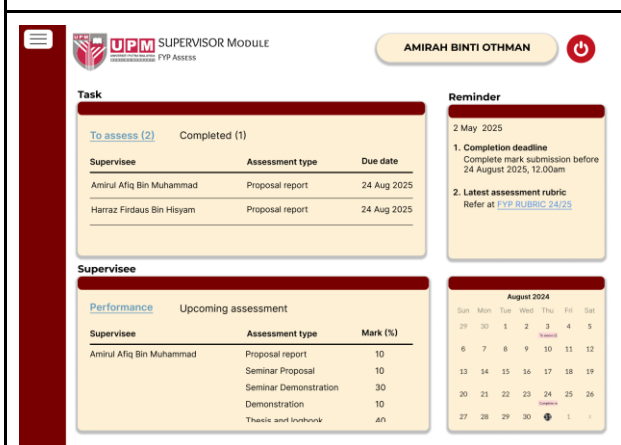
UPM ASSESSOR MODULE
AMIRAH BINTI OTHMAN

Task
To assess (2) Completed (1)
3 Jul 2025, Thu

Student	Time	Venue	Assessors
Siti Athirah Binti Othman	9.00 am	Bilik Kullah A	
Atiya Aisyah Binti Alman	9.30 am	Bilik Kullah A	

Reminder
2 August 2025
1. Completion deadline Complete mark submission before 24 Jul 2025, 12.00am
2. Latest assessment rubric Refer at [FYP Rubric 24/25](#)

Figure 3. Assessor dashboard



UPM SUPERVISOR MODULE
AMIRAH BINTI OTHMAN

Task
To assess (2) Completed (1)

Supervisee	Assessment type	Due date
Aminul Afiq Bin Muhammad	Proposal report	24 Aug 2025
Harraz Firdaus Bin Hisyam	Proposal report	24 Aug 2025

Supervisee
Performance Upcoming assessment

Supervisee	Assessment type	Mark (%)
Aminul Afiq Bin Muhammad	Proposal report	10
	Seminar Proposal	10
	Seminar Demonstration	30
	Demonstration	10
	Thesis and bookb	40

Reminder
2 May 2025
1. Completion deadline Complete mark submission before 24 August 2025, 12.00am
2. Latest assessment rubric Refer at [FYP Rubric 24/25](#)

Figure 4. Supervisor dashboard



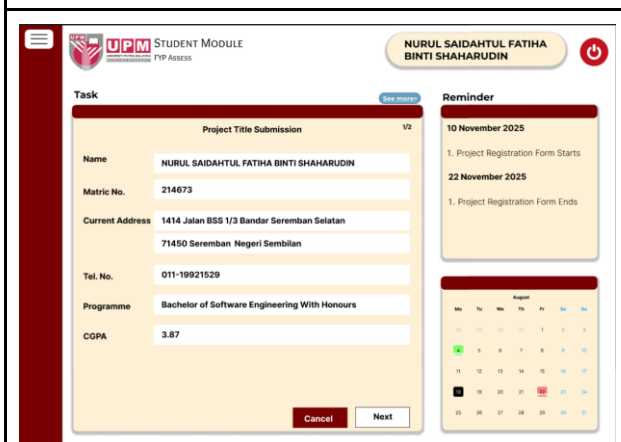
UPM STUDENT MODULE
NURUL SAIDAHTUL FATIHA BINTI SHAHARUDIN

Task
Student CGPA and Resume Submission

Name: NURUL SAIDAHTUL FATIHA BINTI SHAHARUDIN
CGPA: 3.87
Industry: Do you want to do the Final Year Project with any industry? v
Resume: Upload Your File Here

Reminder
10 October 2025
1. Student CGPA and resume submission starts
22 October 2025
1. Student CGPA and resume submission ends

Figure 5. CGPA and Resume Submission



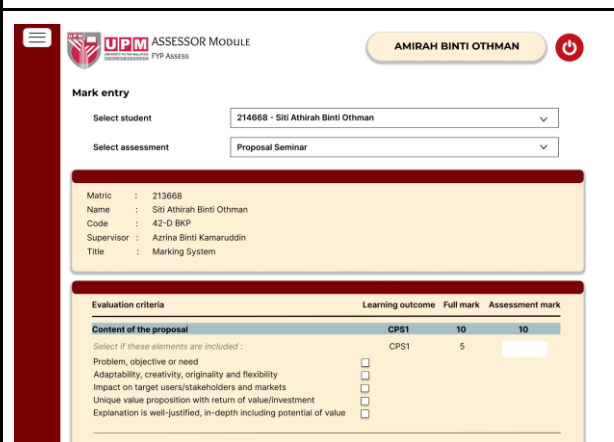
UPM STUDENT MODULE
NURUL SAIDAHTUL FATIHA BINTI SHAHARUDIN

Task
Project Title Submission

Name: NURUL SAIDAHTUL FATIHA BINTI SHAHARUDIN
Matric No.: 214673
Current Address: 1414 Jalan BSS 1/3 Bandar Seremban Selatan, 71450 Seremban Negeri Sembilan
Tel. No.: 011-19921529
Programme: Bachelor of Software Engineering With Honours
CGPA: 3.87

Reminder
10 November 2025
1. Project Registration Form Starts
22 November 2025
1. Project Registration Form Ends

Figure 6. Project Title Submission



UPM ASSESSOR MODULE
AMIRAH BINTI OTHMAN

Mark entry
Select student: 214668 - Siti Athirah Binti Othman
Select assessment: Proposal Seminar

Matric: 213668
Name: Siti Athirah Binti Othman
Code: 42-0 B&P
Supervisor: Azrina Binti Kamaruddin
Title: Marking System

Evaluation criteria	Learning outcome	Full mark	Assessment mark
Content of the proposal	CPS1	10	10
Select if these elements are included:	CPS1	5	
Problem, objective or need			<input type="checkbox"/>
Adaptability, creativity, originality and flexibility			<input type="checkbox"/>
Impact on target users/stakeholders and markets			<input type="checkbox"/>
Unique value proposition with return of value/investment			<input type="checkbox"/>
Explanation is well-justified, in-depth including potential of value			<input type="checkbox"/>

Figure 7. Mark entry

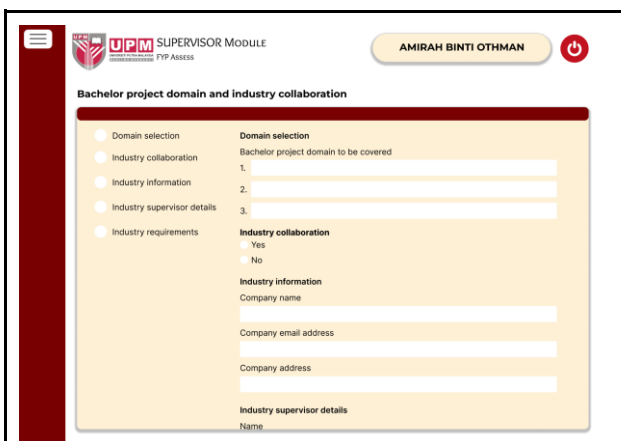
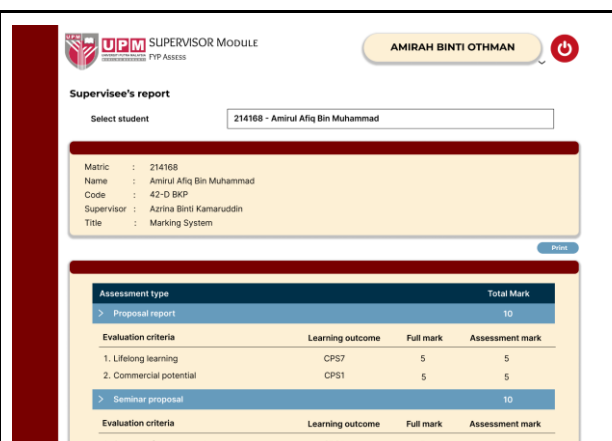
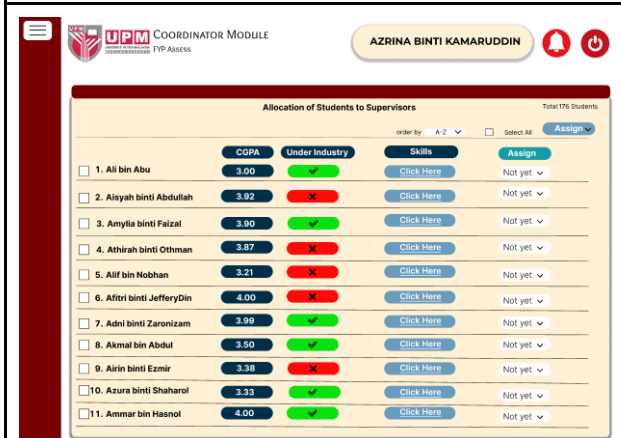


Figure 8. Project domain and industry collaboration form



Assessment type	Total Mark		
Proposal report	10		
Evaluation criteria			
Learning outcome	Full mark	Assessment mark	
1. Lifelong learning	CPS7	5	5
2. Commercial potential	CPS1	5	5
Seminar proposal	10		
Evaluation criteria			
Learning outcome	Full mark	Assessment mark	
1. Content of the proposal	CPS1	5	5

Figure 9. Report generation



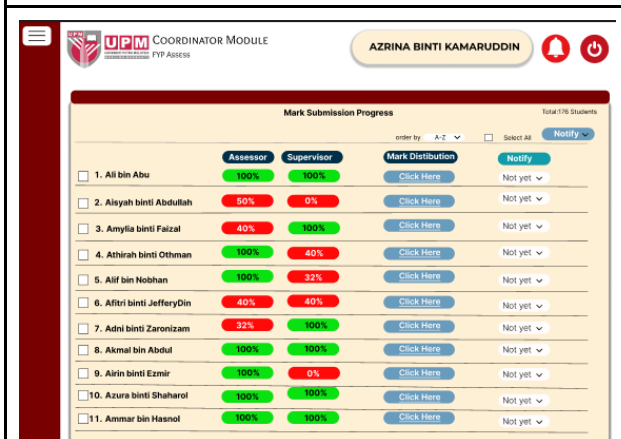
Student Name	CGPA	Under industry	Skills	Assign
1. Ali bin Abu	3.00	✓	Click Here	Not yet
2. Aisyah binti Abdullah	3.92	X	Click Here	Not yet
3. Amylia binti Faizal	3.90	✓	Click Here	Not yet
4. Athirah binti Othman	3.87	X	Click Here	Not yet
5. Alif bin Nobhan	3.21	X	Click Here	Not yet
6. Aftri binti JefferyDin	4.00	X	Click Here	Not yet
7. Adni binti Zaronizam	3.99	✓	Click Here	Not yet
8. Akmal bin Abdul	3.50	✓	Click Here	Not yet
9. Airin binti Ezmir	3.38	X	Click Here	Not yet
10. Azura binti Shaharol	3.33	✓	Click Here	Not yet
11. Ammar bin Hasnol	4.00	✓	Click Here	Not yet

Figure 10. Student Assignment Page



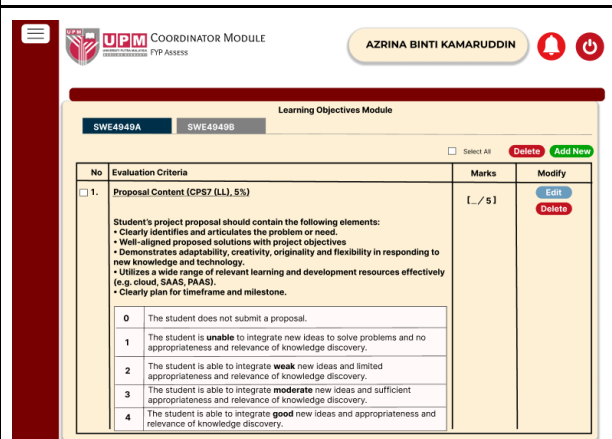
- Dr. Sufri has not submitted marks for Ali bin Abu as an assessor.
- Dr. Kamal has not submitted marks for Nurul Iman binti Azmi as a supervisor.
- Dr. Mazlan has not submitted marks for Nur Aisyah binti Rahman as a supervisor.
- Dr. Salmah has not submitted marks for Mohd Firdaus bin Hassan as a supervisor.
- Dr. Farid has not submitted marks for Muhammad Hafiz bin Ismail as an assessor.
- Dr. Idris has not submitted marks for Alina Sofea binti Khalid as a supervisor.
- Dr. Liyana has not submitted marks for Siti Zulaikha binti Omar as an assessor.
- Dr. Hanim has not submitted marks for Syafiq bin Roslan as a supervisor.
- Razak has not submitted marks for Ahmad Faiz bin Abdullah as an assessor.
- Rahman has not submitted marks for Farhana binti Zainal as a supervisor.
- Hanafi has not submitted marks for Nur Amirah binti Salleh as an assessor.
- Shahrul has not submitted marks for Azlan bin Idris as an assessor.
- Zulkifli has not submitted marks for Amirul Hakim bin Jamal as a supervisor.

Figure 11. Notification Page



Student Name	Assessor	Supervisor	Mark Distribution	Notify
1. Ali bin Abu	100%	100%	Click Here	Not yet
2. Aisyah binti Abdullah	50%	0%	Click Here	Not yet
3. Amylia binti Faizal	40%	100%	Click Here	Not yet
4. Athirah binti Othman	100%	40%	Click Here	Not yet
5. Alif bin Nobhan	100%	32%	Click Here	Not yet
6. Aftri binti JefferyDin	40%	40%	Click Here	Not yet
7. Adni binti Zaronizam	32%	100%	Click Here	Not yet
8. Akmal bin Abdul	100%	100%	Click Here	Not yet
9. Airin binti Ezmir	100%	0%	Click Here	Not yet
10. Azura binti Shaharol	100%	100%	Click Here	Not yet
11. Ammar bin Hasnol	100%	100%	Click Here	Not yet

Figure 12. Mark Submission Progress Page



No	Evaluation Criteria	Marks	Modify
1.	Proposal Content (CPS7 (L1), 5%)	[/ 5]	Edit / Delete
0	The student does not submit a proposal.		
1	The student is unable to integrate new ideas to solve problems and no appropriateness and relevance of knowledge discovery.		
2	The student is able to integrate weak new ideas and limited appropriateness and relevance of knowledge discovery.		
3	The student is able to integrate moderate new ideas and sufficient appropriateness and relevance of knowledge discovery.		
4	The student is able to integrate good new ideas and appropriateness and relevance of knowledge discovery.		

Figure 13. Learning Objective Page

Feedback was collected from the various user's persona including students, supervisors, assessors and coordinators to evaluate the usability and functionality of the version 1

prototype. The comments are summarized as follows: (*S = Student, SV = Supervisor, A = Assessor, C = Coordinator*)

Positive feedback from the formative evaluation session:

1. SV1 stated that the interface layout is clear and organized, providing accessibility by allowing users to easily identify the module and navigate through the menu in the navigation bar.
2. SV2 mentioned that the prototype demonstrates an effective functional flow that aligns with user requirements and addresses the challenges in the current manual assessment process.
3. A1 highlighted that the system allows switching between modules seamlessly without requiring multiple logins, which improves usability

Constructive feedback from the formative evaluation session:

1. S1 suggested standardizing button shapes and interactive elements across all pages to ensure consistency.
2. C1 recommended including brief user instructions or tooltips to assist navigation and prevent confusion.
3. C2 stated that a confirmation pop-up after form submission would provide assurance to the user that the data has been successfully recorded.

FUTURE WORK

In this article we presented the outcome from empathy, define, ideation and prototyping. For the prototyping phase we managed to complete few UI's mock-ups from four user personas. However, the mock-ups need to go through several iterations for further improvement and validation from real users about designing features with the intuitive and usable UI elements, layout and interactions. The next research activities the researcher will conduct requirement and UI's validation with the supervisors, coordinators, students.

CONCLUSION

This paper addresses the issues mainly faced by the coordinators in each department at FCSIT, UPM. One of the issues is that the coordinators need to handle the mark entry process in the system for every single student. Because this process has a high risk of mistakes, it increases the time taken and the workload of the coordinators. This issue is resolved with FYPAAssess as the solution. The solution is developed using a user-centered approach, with a thorough set of interviews done to ensure correct findings and information are gained throughout the development of the solution. This solution supports the fourth Sustainable Development Goal (SDG 4), which focuses on the quality of education.

FYPAAssess accommodates four main types of users, each using the system according to their role, ensuring that the Bachelor Project assessment can be conducted seamlessly and efficiently. With FYPAAssess, supervisors and assessors can input students' marks without performing manual calculations or managing a complex assessment rubric, improving the

efficiency of the assessment process. The coordinators' workload is reduced, as the mark entry process is no longer their responsibility. The system also allows coordinators to track the progress of mark submissions by supervisors and assessors and send notifications to ensure timely submission. Any changes to learning objectives or supervisor assignments for the Bachelor Project course can be updated easily through the system. Additionally, students can modify their Bachelor's Project title directly in the system without repeatedly refilling forms.

ACKNOWLEDGEMENT

We want to thank our participants from the Faculty of Computer Science and Information Technology, UPM, who have been involved in this project.

REFERENCES

- Affandi, R., Bee, T. C., & Baharom, A. (2022). Development of The Mobile Application for Final Year Project (FYP) Management System. *International Journal of Education and Pedagogy*, 4(3), 165-176.
- Alammary, A., Masoud, S., (2025). Towards Smarter Assessments: Enhancing Bloom's Taxonomy Classification with a Bayesian-Optimized Ensemble Model Using Deep Learning and TF-IDF Features. *Electronics*, 14,2312. <https://doi.org/10.3390/electronics14122312>
- CC2020, (2020) *Computing Curricula 2020:Pradigms for Global Computing Education*. Association for Computing Machinery. New York, USA
- Edupuganti, R., (2013). *Online grading system* (Doctoral dissertation, California State University, Sacramento).
- Hironor. W.,(2024) eds., *Guide to the Software Engineering Body of Knowledge (SWEBOK Guide), Version 4.0*, IEEE Computer Society, 2024; www.swebok.org.
- Ismail, S. I., Abdullah, R., Kar, S. a. C., Fadzal, N., Husni, H., & Omar, H. M. (2017). Online Project Evaluation and Supervision System (oPENs) for final year project proposal development process. *IEEE 15th Student Conference on Research and Development* <https://doi.org/10.1109/scored.2017.8305392>
- Kar, S. A. C., Fadzal, N., Ismail, S. I., Abdullah, R., & Omar, H. M. (2017). Integrated supervision and evaluation system for final year project. In *2017, IEEE 15th Student Conference on Research and Development (SCOREd)* (pp. 198-203). IEEE.
- Kusay, M. P. S. (2023, October 10). *ECS Final year project marking system*. <https://ojs.victoria.ac.nz/wfes/article/view/8410>
- Ning, Y. (2025). Integrating Design Thinking into Higher Education: Enhancing Student Creativity and Critical Thinking. *Journal of Higher Education Research*, 5(6), 564-566.
- Simanjuntak, W. M. M. C., & Ghazali, A. (2025). Improving bank digital services using design thinking method. *Eduvest - Journal of Universal Studies*, 5(2), 2172–2184. <https://doi.org/10.59188/eduvest.v5i2.50867>
- Suria, O. (2024). A Statistical analysis of System Usability scale (SUS) evaluations in online learning platform. *Journal of Information Systems and Informatics*, 6(2), 992–1007. <https://doi.org/10.51519/journalisi.v6i2.750>
- Thompson, M., & Ahn, B. (2012). The Development of an Online Grading System for Distributed Grading in a Large First-year Project-based Design Course. *Proceedings of the 119th ASEE Annual Conference & Exposition*, 25.11.1-25.11.14. <https://doi.org/10.18260/1-2--20767>

AI-Enhanced Hybrid Metaverse Framework for Multiplayer Interactive and Gamified Virtual Campus Tours

Sagar Prema Rajan, Siti Azreena Mubin, Vinesh Thiruchelvam
Asia Pacific University of Technology & Innovation, Malaysia

ABSTRACT

This research presents an AI-driven Hybrid Metaverse Framework (HMF) for immersive, interactive, and simultaneous virtual campus visits and tours. The HMF allows prospective students, academic staff, and other stakeholders to form and engage with a university's submerging virtual campus. The HMF framework integrates AI, Virtual Reality (VR), and gamification prompts with Adaptive, Context-Sensitive Non-Playable Characters (NPCs) to facilitate participant's onboarding to university environments, while simultaneously exploring collaboratively and in real time. In a pilot exploration study with ten participants (7 students and 3 VR experts), the framework increased levels of engagement, social presence, completion of tasks, and knowledge retention levels compared to single-user virtual campus visits and tours. A combination of descriptive statistics and paired t-tests confirmed there were statistically meaningful practical differences in user experience. Further, qualitative thematic analysis confirmed strong levels of perceived immersion, collaborative learning, and personal space for exploration. Therefore, the research demonstrates the HMF is an innovative method for the purpose of outreach, orientation and recruitment for the university posed, that is scalable and adaptable.

Keywords: Artificial Intelligence, Metaverse, Virtual Campus Tours, Multiplayer Learning Environments, Gamification, Hybrid Framework, Immersive Education, AI-driven NPCs

INTRODUCTION

The growing convergence of Virtual Reality (VR), Artificial Intelligence (AI), and 3D immersive media have facilitated the development of the metaverse as an engaging educational medium (Mystakidis, 2022). Colleges and universities' interest in this digital space comes from the opportunity to provide an exciting and interactive campus experience to prospective students, faculty, and visitors that goes beyond traditional constraints of location, cost, and the lack of personalized and collaborative exploration of campuses (Zhang & Lee, 2022).

Early virtual campus tour solutions typically adopted single-user static view formats, offering little more than passive walkthroughs. These early attempts lacked adaptive intelligence, meaningful social interaction, and gamification, which are key elements for creating engaging and memorable experiences (Díaz & Saldaña, 2023; Bailenson et al., 2021). With the advent of AI-driven natural language processing (Zhao et al., 2023; Cheng & Tsai, 2023) and multiplayer networking interactions (Garcia & Huang, 2023), there is now an opportunity to

offer virtual campus tours as interactive, engaging, collaborative, and social experiences, enhancing both accessibility and immersion.

To address this gap, this paper introduces a Hybrid Metaverse Framework (HMF) that integrates AI-enabled non-playable characters (NPCs), multiplayer collaboration, and gamification to create a distinct and scalable virtual campus tour experience. The framework is customizable to various institutions and supports both immersive VR and desktop-based access, allowing flexible participation.

This work also aligns with the United Nations Sustainable Development Goals (SDGs) particularly SDG 4 (Quality Education) and SDG 9 (Industry, Innovation and Infrastructure) by promoting accessible, inclusive, and innovative digital learning experiences.

Paper Objectives:

1. To design and implement a gamified virtual campus tour integrating AI-driven interactions and multiplayer collaboration within a hybrid metaverse framework.
2. To compare the effectiveness of single-player and multiplayer modes in enhancing engagement, realism, and collaborative learning experiences.
3. To evaluate how AI-enhanced gameplay elements influence user interaction, immersion, and perceived educational value in virtual campus environments.

BACKGROUND

The word metaverse describes a connected system of immersive, three-dimensional virtual spaces in which users share the ability to interact with each other as well as with digital objects in real-time. In an educational context, the metaverse provides opportunities for remote learners to participate, collaborative or shared learning experiences, and experiences that allow for expeditions into the unknown without the restrictions of physical space (Mystakidis, 2022). The role of Virtual Reality (VR) technology allows for realistic and navigable environments, while Artificial Intelligence (AI) allows for responsive and individualized interactions with meaningful adaptation to each learner's needs (Zhao et al., 2023).

Recent literature has demonstrated evidence of pedagogy and possibilities that can be derived from combining immersive environments with intelligent systems. Zhang and Lee (2022) demonstrated the advantages of hybridized models of learning within higher education combining physical and virtual elements that offer flexibility and scalability. The social presence, collaboration, and trust captured through avatar-based communication in multiplayer VR environments offer leverage for learning and student engagement (Bailenson et al., 2021).

Gamification is the application of game elements in a context that is not a game. Gamification has been studied predominantly as a means of increasing motivation and retention in educational settings. Evidence from Díaz and Saldaña (2023) showed that metaverses created

gamified activities which are immersive, such as quests and achievements, lead to better improvement in learning outcomes and learner satisfaction in study participants. Similarly, Papachristos et al. (2022) found immersive VR and multi-user experiences supported collaborative learner-to-learner interaction which, in turn, aided the deeper retention of knowledge.

AI is another facet that may contribute to enhancing learner experiences within the metaverse learning environment. Cheng and Tsai (2023) highlighted AI tutors or toolkits use learning data to adapt learning content delivery to the point of progress for that learner and Garcia and Huang (2023) showed that facilitation, or moderation, of social dynamics in multiplayer classrooms through AI, immediately improved the social dynamics of the group. Lee and Kim (2024) further advanced this understanding of AI when they created adaptive gamified systems for education, where feedback was provided to personalize challenges for each learner as well as their own output from reflection as feedback.

While advancements have occurred in using, linking, and integrating AI, multiplayer networking, and gamification as part of this work, there is still no cohesive, scalable framework for virtual campus tours that combine elements of AI, networking, and gamification together. Typically, these elements are treated separately, which limits their potential individually or as combined social-rich, adaptive engaging experiences. This paper introduces the Hybrid Metaverse Framework (HMF), which is a system that brings these together so as to let institutions examine recruiting, orientation, and outreach and customize it for their particular use.

METHODOLOGY

This study adopted a mixed-methods pilot exploratory design that combined system development with both quantitative and qualitative evaluation. The Hybrid Metaverse Framework (HMF) was conceptualized to integrate VR, AI, and multiplayer networking to create immersive, gamified, and collaborative campus tours.

Research Questions

1. Does the AI-enhanced multiplayer HMF increase user engagement and immersion compared to a single-user tour?
2. How do users perceive social presence and collaborative learning in the HMF environment?
3. What qualitative experiences emerge from interacting with AI-driven NPCs in multiplayer settings?

System Architecture

The HMF consists of four layers:

- (1) *Client Interface Layer* supporting VR headsets and desktops.

(2) *Interaction Layer* using Photon Unity Networking for synchronous low-latency communication.

(3) *AI Interaction Layer* enabling NPC activation through a Natural Language Processing engine; and

(4) *Content Management Layer* allowing institutions to upload and configure interactive objects and gamified content. NPCs provided context-aware responses based on user location, task, and session history. Multiplayer functionality synchronized movement, voice, and collaboration, while scavenger hunts, mini-games, and badges added gamification (Díaz & Saldaña, 2023; Bailenson et al., 2021).

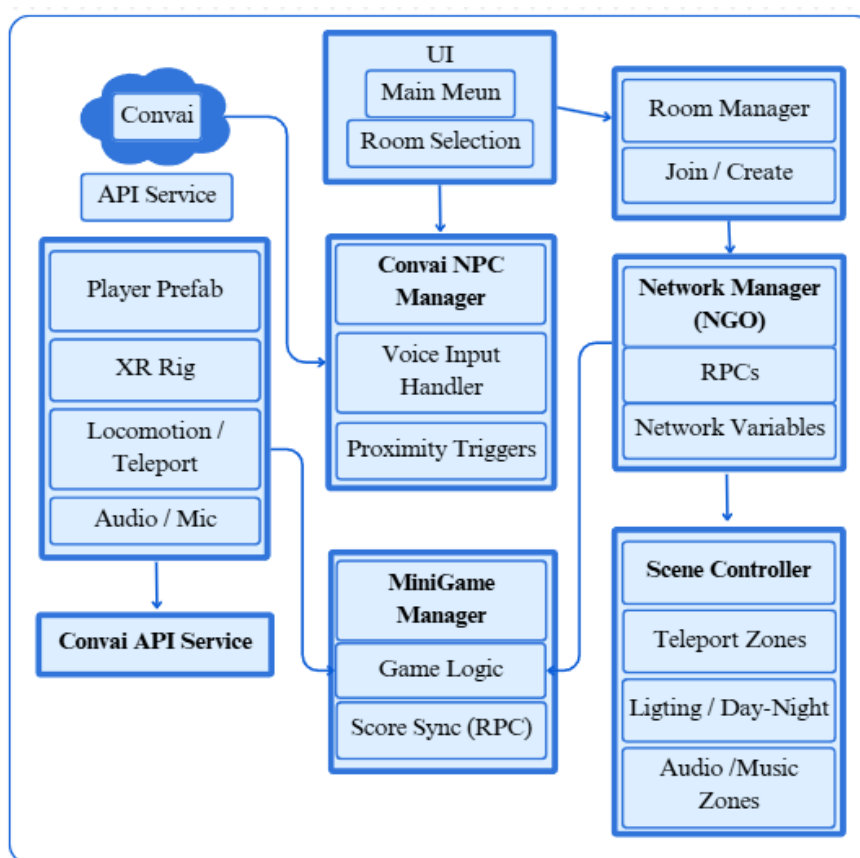


Figure 1: System Architecture

Participants and Ethics

Ten participants (7 students, 3 VR experts) were recruited using convenience sampling. Ethical clearance was obtained, informed consent secured, and anonymity ensured.

Evaluation Approach

Each participant experienced a single-user traditional tour and a multiplayer AI-enhanced tour. Data collection included surveys on engagement and immersion, observation logs on interaction behaviors, and semi-structured interviews on social presence, usability, and AI responsiveness.

Data Analysis

Quantitative data was analyzed using descriptive statistics and paired t-tests to compare conditions. Qualitative interview transcripts were coded and analyzed thematically using NVivo to identify recurring patterns. This approach allowed systematic triangulation of quantitative outcomes with qualitative insights. The methodology builds on prior research showing that AI-enhanced multiplayer VR with gamification increases engagement and retention, offering a scalable model for virtual campus experiences.

RESULTS AND FINDINGS

The evaluation compared the multiplayer AI-enhanced HMF with conventional single-user virtual campus tour. The review was relevant to three main areas, while a range of other themes were considered which are not included in this report. The three areas of focus included whether the tour positively influenced the participant's understanding of the campus in a meaningful way, whether it allowed for real-time communication and playful interaction and whether the VR experience felt real and immersive.

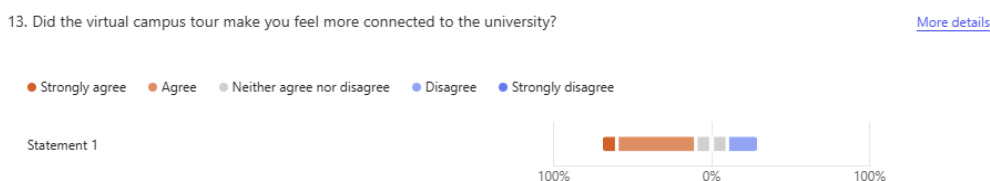


Figure 1. Survey Question

Table 1. Survey results indicated a clear advantage for the multiplayer AI-enhanced tour

Survey Dimension	Single-User Tour	Multiplayer AI-Enhanced Tour	t (9)	p value
Perceived Campus Understanding	3.6	4.5	3.12	0.012
Ability to Talk and Interact	2.8	4.7	4.56	0.001
VR Realism and Immersion	3.9	4.6	2.41	0.038

Note: All scores were measured on a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree). The significance was established with paired t-tests.

The results showed statistically significant differences in all three dimensions. Users all had more engagement, greater communication and interaction, and a better sense of immersion with multiplayer AI-enhanced tours compared with single-user conditions.

Perceived Understanding of Campus: $t(9) = 3.12, p = 0.012$

Speaking and Interacting: $t(9) = 4.56, p = 0.001$

Realism and Immersion in VR: $t(9) = 2.41, p = 0.038$

While there was a small sample size, the effect sizes were large, signaling an important practical difference in engagement and immersion for the multiplayer AI-enhanced tours.

Qualitative Insights

Responses to the interview showed the participants appreciated a combination of AI direction, multiplayer collaboration, and immersive exploration opportunities through the Hybrid Metaverse Framework (HMF). Analysis of the interviews through NVivo revealed three major themes:

1. Seamless Engagement of AI and Multiplayer Engagement

Participants focused on the combination of real-time social engagement along with AI-directed non-playable characters (NPC), allowing them to find a balance between becoming supported navigators and moving into their own independent explorations. This hybrid engagement enhanced their experiences overall and enabled significant collaboration with others.

Student 2: "The AI guide provided me with valuable information right when I needed it, along with being able to ask my peers questions and explore together with few restrictions."

2. Flexible and Inclusive Access Across Multiple Platforms

Participants appreciated the exposure to the metaverse via both some form of VR headset and alternative access while on a computer desktop, which emphasizes the inclusive potential of the hybrid access approach. This simultaneously was considered crucial for broader scalability and access for users.

Expert 1: "It's so rare to see a framework that connects so seamlessly together and work together in the context of immersive and non-immersive frameworks; this makes it accessible and able to be scaled to new users."

3. Game-like Exploration with Contextual Learning

Participants viewed their experience as game-like, with exploring being rewarded with contextual learning opportunities from the AI system. Because of this design, learners were motivated to actively participate and were more engaged and connected to the activity.

Student 4: "Like a game where the rewards are knowledge you are always surprised with what you have uncovered and learned with the system."

The findings show interactions driven by AI, collaborations between students in a multiplayer environment, and gamified learning aspects work together mutually to improve social presence, engagement, and knowledge retention in virtual campus tours.

Interpretation of Results

The results indicate that the HMF's integration of AI NPCs, multiplayer features, and high-fidelity VR environments not only supported understanding of the campus but also encouraged social interaction and immersion, which are all known to help engagement with virtual learning (Bailenson et al., 2021; Díaz & Saldaña, 2023). The participants' freedom to explore the campus as they wished, instead of completing a prescribed set of tasks, supported them in creating their own meaningful experience of the virtual campus.

DISCUSSION AND FUTURE WORK

The findings of this exploratory study indicate that the hybrid metaverse framework (HMF), with its use of AI-powered non-player characters (NPCs), multiplayer cooperative activities, and gamification features, is able to improve student engagement, immersive learning atmosphere, and social presence during a virtual campus experience. These findings are in line with previous work suggesting that multiplayer VR settings engage students more and support knowledge retention (Bailenson et al., 2021; Díaz & Saldaña, 2023). While the sample was small, the results provided some initial evidence for the scalability, flexibility, and educational effectiveness of the framework.

As immersive technologies, AI, and connectivity technology improve, the HMF will have new aesthetic and functional possibilities. With inexpensive hardware and faster internet speeds, it may be possible to deliver low-latency, large-scale integrated multiplayer campus experiences at low visual fidelity. One emerging design principle involves adaptive

learning analytics in which AI systems are observing participant behaviors, preferences, and progress in real time. These features could facilitate real-time adaptations of content complexity, modes of learning, and pathways for exploration for each of the participants.

Future versions of the framework could implement large language models to enhance NPC interactions and capabilities, making them more human-like, contextualized, and easy to communicate with. Optimizing for cross-device compatibility to increase accessibility will be important, like potential integration within LMS platforms, that could facilitate and maintain schools' abilities to use learning outcomes, engagement metrics, and competencies.

Future Research Suggestions

- Study the long-term outcomes of AI-infused metaverse campus tours for student recruitment, retention, and branding purposes.
- Analyze the impact of gamification across settings of varied complexity and culture.
- Compare the experiences of multiplayer tours led by AI as compared to human-led tours with respect to knowledge transfer and student engagement.
- Explore accessibility features such as text-to-speech, haptic feedback, and AI guides across multiple languages.

In the future, the HMF could help support the design of fully immersive smart campuses that connect virtual tours to lectures, labs, and student engagement. This new paradigm could reshape the thoughtfully interactive experience that prospective students have when learning about and interacting with institutions of higher education.

CONCLUSION

This paper has described the design, development, and evaluation of a Hybrid Metaverse Framework (HMF) that uses artificial intelligence, multiplayer interaction, and gamified exploration to create more engaging and immersive virtual campus tours. By combining adaptive AI-guided navigation, synchronous social interaction, and cross-platform access, the framework overcomes constraints of traditional single-user virtual campus tours, creating a more immersive, engaging, and informative experience.

The findings of this study demonstrate that the HMF meaningfully enhances participants understanding of campus facilities, enables real-time communication and collaborative exploration, and enhanced perceived realism of the virtual environment. While the small sample limited the availability of statistical significance, both quantitative trends and qualitative feedback suggest that the hybrid design paradigm has great potential for educational institutions looking to maximize access and engagement.

The principles exhibited by the HMF can be used in contexts beyond the campus tours it provides, noting many more educational and training contexts, are possible in the metaverse. By leveraging AI, gamification, and scalable multiplayer immersion, institutions are able to provide new forms of analogous learning experiences for in-person classrooms, virtual environments, and hybrid learning spaces. This writing expands the literature base

of immersive educational technology and provides a framework for further progress towards a fully integrated AI driven virtual learning ecosystem.

ACKNOWLEDGMENT

The author wishes to sincerely thank the students and VR experts who devoted their time, participation, and thoughtfulness to help evaluate the Hybrid Metaverse Framework. The author wishes to also thank the faculty and technical staff who, with graciousness, provided access to institutional resources and engaged in the virtual campus development process. The author would also like to acknowledge the academic mentors and peers whose mentoring and inputs contributed to shaping the direction of this research. Appreciate and respect is also extended to the 7th National Symposium on Human Computer Interaction 2025 (FUSION2025) for providing an opportunity to share findings and requests for feedback which reinforced the final work.

REFERENCES

- Bailenson, J. N., Blascovich, J., & Beall, A. C. (2021). Avatars and collaborative virtual environments: Design and evaluation of social interaction in immersive VR. *Presence: Teleoperators and Virtual Environments*, 30(2), 123–140. https://doi.org/10.1162/pres_a_00321
- Cheng, Y., & Tsai, C.-C. (2023). Artificial intelligence in education: Adaptive learning and personalization in immersive environments. *Computers & Education*, 197, 104743. <https://doi.org/10.1016/j.compedu.2023.104743>
- Díaz, J., & Saldaña, M. (2023). Gamification in the metaverse: Enhancing student engagement and learning outcomes in higher education. *Education and Information Technologies*, 28(4), 4583–4602. <https://doi.org/10.1007/s10639-023-11488-1>
- Garcia, P., & Huang, L. (2023). AI moderation in multiplayer learning environments: Strategies for maintaining engagement and inclusivity. *British Journal of Educational Technology*, 54(5), 1432–1450. <https://doi.org/10.1111/bjet.13342>
- Lee, S., & Kim, H. (2024). Adaptive gamified learning systems in virtual reality environments: Design and evaluation. *Interactive Learning Environments*, 32(1), 56–74. <https://doi.org/10.1080/10494820.2022.2064321>
- Lu, M., & Xu, J. (2023). Content management frameworks for scalable educational metaverse platforms. *Journal of Educational Technology Systems*, 52(2), 257–278. <https://doi.org/10.1177/00472395231158720>
- Mystakidis, S. (2022). Metaverse. *Encyclopedia*, 2(1), 486–497. <https://doi.org/10.3390/encyclopedia2010031>
- Papachristos, N. M., Vrellis, I., & Mikropoulos, T. A. (2022). Learning in multiplayer immersive virtual environments: Collaborative problem solving and knowledge retention. *Educational Technology Research and Development*, 70(3), 1235–1256. <https://doi.org/10.1007/s11423-021-10043-9>
- Zhang, X., & Lee, M. (2022). Hybrid learning in higher education: Opportunities and challenges in the post-pandemic era. *Journal of Computer Assisted Learning*, 38(6), 1505–1518. <https://doi.org/10.1111/jcal.12676>
- Zhao, Q., Li, X., & Wang, Y. (2023). Intelligent conversational agents for education: Applications in virtual and augmented reality. *Computers in Human Behavior*, 139, 107523. <https://doi.org/10.1016/j.chb.2022.107523>

Towards Considerate AI: Building a Fair and Polite Chatbot for Customer Service

Looi Pei Yi¹, Ng Jin Jin², Li ZhiHan³, Raja Jamilah Raja Yusof⁴, Viktor Kaptelinin⁵
^{1,2,3,4}Universiti Malaya, Malaysia
⁵Umeå Universitet, Sweden

ABSTRACT

This study presents a customer service chatbot designed to communicate fairly and politely using a large language model (LLM). While many chatbots prioritize efficiency, they often neglect tone and bias, which affect user trust and satisfaction. The chatbot was developed using Flowise with Retrieval-Augmented Generation (RAG) and Pinecone, incorporating principles from politeness theory and fairness research. The system supports three service scenarios: product recommendations, order tracking, and FAQs. Evaluation through user feedback and heuristic assessment indicates that respectful and emotionally aware responses improve user perception and interaction quality. The findings highlight the importance of ethical language design in developing socially intelligent conversational agents.

Keywords: Politeness, Fairness, Chatbot, LLMs, Customer Service, Heuristic Evaluation, RAG, Flowise, Pinecone

1.0 INTRODUCTION

1.1 Background and Motivation

Conversational agents, particularly chatbots, are widely used in customer service, education, and healthcare to provide scalable and continuous support. Powered by natural language processing (NLP) and large language models (LLMs), these systems simulate human-like interaction while improving accessibility and reducing operational workload. However, technical capability alone is insufficient. Users increasingly expect AI systems to demonstrate social intelligence, particularly fairness and politeness. Fairness refers to equitable and unbiased treatment, while politeness involves respectful communication. These qualities strongly influence user trust and satisfaction. Despite advances in LLMs, many chatbots still produce responses perceived as robotic or insensitive, negatively affecting user experience (Liebrecht, Sander, & van Hooijdonk, 2021). Therefore, integrating fairness and politeness into chatbot design remains an important research challenge.

1.2 Objectives

This study aims to design and evaluate a customer service chatbot that incorporates fairness and politeness in its communication. First, the study examines user perceptions of fair and polite chatbot interactions in e-commerce contexts. Second, these insights are translated into practical design principles for chatbot dialogue and interface development. A generative AI model (GPT-4) is used to implement these principles and enable context-aware responses. Finally, the system is evaluated using heuristic assessment and user feedback to determine its effectiveness in supporting fair and polite communication.

2.0 LITERATURE REVIEW

2.1 Politeness and Fairness

Politeness and fairness are key principles in human communication and increasingly important in chatbot design. Politeness refers to respectful language, while fairness ensures equitable and unbiased treatment. Together they influence user trust and engagement with conversational systems.

Politeness theory explains how individuals manage two types of “face”: positive face (the desire to be appreciated) and negative face (the desire for autonomy) (Brown & Levinson, 1987). Strategies such as positive politeness, negative politeness, and indirect communication reduce face-threatening acts during interaction (Al-Duleimi, 2016; Shahrokhi & Bidabadi, 2013). Rapport management theory further emphasizes maintaining respectful relationships in communication (Spencer-Oatey, 2000).

Fairness focuses on ethical and consistent treatment. According to Fairness Theory, individuals evaluate fairness based on intentions, consequences, and accountability (Folger & Cropanzano, 2001). Consistent rule application is also essential for perceived fairness (McMahon, 2016). Violations of fairness expectations may lead to dissatisfaction or distrust (Kim, 2011). Integrating politeness and fairness, therefore, helps create trustworthy and socially intelligent chatbot interactions.

Table 1. Principles of Politeness and Fairness in Chatbot Design

Principle	Description	Example / Application	Key References
Emotionally Intelligent Communication	Chatbot acknowledges user emotions and responds empathetically.	“I understand your concern. Let me help check that for you.”	Spencer-Oatey (2000); Song et al. (2023); Brown & Levinson (1987)
Face-Saving Communication	Uses indirect and respectful phrasing to reduce social pressure.	“Would you like me to prepare a list of options?”	Brown & Levinson (1987); Shahrokhi & Bidabadi (2013)
Sociability and Trust Building	Friendly language improves engagement and human-likeness.	“You’re welcome! Let me know if you need anything else.”	Brummernhenrich et al. (2025); Song et al. (2023)
Appreciation-Based Responses	Express gratitude rather than excessive apologies.	“Thank you for your patience.”	Song et al. (2023)
Consistent and Fair Treatment	Responses follow consistent rules for all users.	Avoid biased recommendations based on user profile.	Folger & Cropanzano (2001); McMahon (2016)
Transparency and Accountability	Chatbot explains decisions and system limitations.	Explaining refund policies or recommendation logic.	Diakopoulos (2016); Assen (2024)
Bias Mitigation	Avoid discriminatory language and ensure inclusive responses.	Use diverse training data and neutral phrasing.	Xue et al. (2024); Mehrabi et al. (2021)

2.2 Use of LLMs in Conversational Interface

Large Language Models (LLMs) are AI systems capable of understanding and generating human language. Built on transformer architectures and trained on large text datasets, models such as GPT-3, GPT-4, and T5-XL support tasks including text generation, question answering, translation, and summarization (Zhao et al., 2024). A key advantage of LLMs is their ability to maintain context during multi-turn conversations, allowing chatbots to produce coherent responses aligned with user intent. In customer service environments, LLM-based chatbots can automate routine inquiries, handle large volumes of requests, and provide continuous support (Uzoka et al., 2024).

However, LLM performance varies across tasks and datasets. Larger models do not always guarantee better results, and outcomes depend on fine-tuning and task-specific optimization. General-purpose models may also lack evaluation mechanisms tailored to customer service contexts (Li et al., 2024).

2.3 Heuristic Evaluation in Chatbot Design

Heuristic evaluation is widely used to assess chatbot usability and interaction quality. Unlike graphical interfaces, conversational agents involve dynamic, turn-based interactions requiring evaluation of dialogue flow, intent recognition, and language clarity.

Researchers have proposed chatbot-specific heuristics addressing conversation coherence, error recovery, and system transparency (Langevin et al., 2021). Evaluation frameworks such as those proposed by Essop et al. (2023) measure chatbot performance across dimensions including accuracy, dialogue coherence, and responsiveness. Sugisaki and Bleiker (2020) also combine heuristic evaluation with ISO 9241-11 usability standards to integrate subjective and objective performance metrics. Applications across domains such as business rule validation, cybersecurity training, and educational chatbots show that heuristic evaluation improves usability, task completion, and engagement (Bockhorst et al., 2019; Pears et al., 2021; Pessoa et al., 2024).

2.4 Research Gap

Although conversational agents are widely used, most systems prioritize task efficiency and language fluency while overlooking social interaction factors such as fairness and politeness. Existing studies acknowledge their importance but rarely provide systematic methods for translating these concepts into practical chatbot design strategies. Furthermore, while LLMs such as GPT-4 generate fluent responses, they lack mechanisms for handling biased or inappropriate user input without additional design intervention. Increasing model size alone does not ensure socially responsible behavior. Finally, current heuristic evaluation methods mainly focus on usability rather than social qualities such as emotional awareness or bias mitigation. Standardized evaluation criteria for fairness and politeness remain limited. To address these gaps, this study develops a generative AI chatbot that integrates fairness and politeness into dialogue design and evaluates these principles using a heuristic framework that combines usability and socio-ethical criteria.

3.0 METHODOLOGY

3.1 Development Approach

This study adopts a User-Centered Design (UCD) approach to develop a fair and polite customer service chatbot. The process consists of five stages: Define, Research, Ideate, Design, and Test.

The Define stage establishes the chatbot's purpose of supporting product recommendations, order tracking, and FAQs while maintaining respectful interaction. The Research stage collects user insights through interviews and surveys. During Ideate, design requirements such as unbiased recommendations and polite communication are identified. The Design stage develops conversation scripts and interaction flows. Finally, the Test stage evaluates the chatbot through user testing and heuristic analysis.

3.2 LLMs Tools

The chatbot was developed using Flowise, an open-source platform for building LLM applications through a visual interface. Flowise allows integration of GPT-based models with minimal programming. The system uses Retrieval-Augmented Generation (RAG), which combines

information retrieval with text generation. This approach enables the chatbot to retrieve relevant knowledge before generating responses, improving accuracy and reducing hallucination (Lee et al., 2024). Pinecone is used as a vector database to store semantic embeddings. By converting knowledge into vector representations, Pinecone enables efficient similarity search and retrieval of relevant information during conversations (Xie et al., 2023).

3.3 Conversation Design

Conversation design translates fairness and politeness principles into chatbot dialogue behavior. Guided by Politeness Theory (Brown & Levinson) and Fairness Theory (Folger & Cropanzano), dialogue templates were created for common customer service scenarios. Politeness strategies include appreciation-based responses and optional phrasing to reduce face-threatening interactions. Fairness is maintained through consistent responses and avoidance of biased language. When encountering impolite user input, the chatbot applies deflection strategies to maintain respectful communication.

Dialogue templates were refined iteratively and adapted into prompt structures compatible with the LLM.

3.4 Interaction Scenarios and Logic

Three interaction contexts were defined: product recommendations, order tracking, and FAQs. User inputs were categorized into four types based on politeness and fairness: Polite & Fair, Polite & Unfair, Impolite & Fair, and Impolite & Unfair. Scenario-based interactions simulated realistic customer requests, including challenging cases such as rude or unrealistic demands.

When a user message is received, the system first classifies the request by context and then evaluates the level of politeness and fairness. Based on this classification, the chatbot selects an appropriate dialogue strategy that applies politeness techniques and fairness principles to generate a suitable response.

3.5 UI & System Design

The system follows a modular structure similar to Flowise workflows. User input is processed through intent and tone classification modules before being sent to the LLM. Conversation memory maintains interaction history to support multi-turn dialogue. The system integrates GPT-4.1 with Redis for conversation memory and Pinecone for vector retrieval. Product information, order data, and FAQ content are stored as embeddings to enable semantic search. Prompt design ensures that responses remain polite, transparent, and fair. The chatbot interface provides a simple conversational environment where users can submit queries and receive responses in real time.

3.6 User Testing

Three survey instruments were used to evaluate chatbot performance. The first survey assessed fairness and politeness in conversational behavior using items from the Godspeed Questionnaire Series, measuring perceptions such as anthropomorphism, likability, and perceived intelligence. Two additional surveys were used for model comparison, evaluating chatbot responses generated from different document sizes and complexities. Participants completed predefined tasks and rated the chatbot based on clarity, helpfulness, tone, fairness, and politeness. A comparative evaluation was also conducted across three large language models—LLaMA, DeepSeek, and ChatGPT. Participants interacted with each model and evaluated response quality, tone, and usefulness, while developers assessed technical factors such as response latency and operational costs.

3.7 Heuristic Evaluation Setup

A heuristic evaluation framework focusing on fairness and politeness was developed. The criteria included respectful responses, non-imposing language, trust-building communication, appreciation-based phrasing, consistent reasoning, equal treatment, contextual awareness, and transparency.

Two evaluators a senior full-stack developer and an algorithm engineer—interacted with the chatbot using typical tasks such as product inquiries, order tracking, and FAQ navigation. They assessed both response quality and interaction tone based on the defined heuristics. For each task, evaluators identified heuristic violations, recorded examples, rated severity on a 0–4 scale (0 = no issue, 4 = critical issue), and suggested improvements where necessary.

4.0 RESULTS

4.1 Chatbot Functionality

The chatbot supports product recommendations, order tracking, and FAQ responses. Using GPT-4.1 with a retrieval framework, the system retrieves knowledge from a vector database and generates context-aware replies. For product recommendations, the chatbot collects preferences such as product type and price range. Order tracking allows users to check delivery status via order ID, while FAQ queries retrieve policy information. Conversation memory supports coherent multi-turn dialogue.

4.2 Politeness and Fairness Evaluation Outcomes

Seven participants evaluated the chatbot using the task list. All participants rated the chatbot positively for politeness and fairness, selecting either “Agree” or “Strongly Agree.” These results suggest that the chatbot maintained a respectful and professional tone during interactions.

4.3 Godspeed Questionnaire Results

Godspeed results show positive user perceptions across dimensions such as friendliness, responsiveness, competence, and intelligence. Most participants rated the chatbot between 4 and 5, indicating that the system was perceived as engaging and human-like.

4.4 Heuristic Evaluation Findings

The evaluation showed strong performance in politeness criteria, particularly respectful language and trust-building communication. However, improvements were identified in fairness aspects such as recommendation explanations and contextual personalization. Recommended improvements include providing clearer recommendation reasoning, collecting user preferences before suggestions, maintaining soft boundaries when users are impolite, and improving contextual personalization.

4.5 Comparison Results

Responses were compared across GPT-4, LLaMA, and DeepSeek. GPT-4 produced the most context-aware and structured responses, offering detailed explanations and proactive assistance. LLaMA generated consistent but less personalized responses, while DeepSeek produced polite but simpler replies. Although LLaMA and DeepSeek offer lower operational costs, GPT-4 provides the highest conversational quality and contextual understanding.

5.0 DISCUSSION

5.1 Interpretation of Findings

The findings show that integrating fairness and politeness improves user trust, satisfaction, and interaction quality. The chatbot maintained respectful communication even when users were impolite or made unrealistic requests. User evaluations and Godspeed results indicate positive perceptions of friendliness, responsiveness, and intelligence. Comparative testing also shows that GPT-4 performs better than alternative models in maintaining context-aware and polite responses. These findings highlight the importance of combining ethical design principles with LLM capabilities when developing socially intelligent conversational systems.

5.2 Limitations

First, the chatbot was implemented as a prototype using a text-based interface, which limited the ability to incorporate multimodal signals such as voice tone or facial expressions that might enhance emotional understanding. Second, although GPT-4 demonstrates strong performance, its responses still depend heavily on prompt design and training data. In some cases, responses required manual tuning to maintain fairness and politeness across different scenarios. Third, the user evaluation involved a relatively small participant group, and most interactions were scenario-based rather than spontaneous. As a result, the findings should be interpreted as preliminary and may not fully represent broader user populations.

6.0 CONCLUSION

This study developed a customer service chatbot incorporating fairness and politeness as core interaction principles. By combining ethical communication theories, user-centered design, and large language models, the system produces respectful and context-aware responses across different scenarios. Evaluation results show that integrating fairness and politeness improves user trust and interaction quality. Future work may extend the system with sentiment analysis, multilingual support, and real-world deployment with larger user populations.

7.0 REFERENCES

- Al-Duleimi, H. Y. (2016). *A critical review of prominent theories of politeness*. *Advances in Language and Literary Studies*, 7(6). <https://doi.org/10.7575/aiac.all.v.7n.6p.262>
- Assen, L. J. (2024). *Quality of interaction with chatbots as a function of fairness, appearance and congruence with participants' declared sex* [Bachelor's thesis, University of Twente]. <https://purl.utwente.nl/essays/100688>
- Binns, R. (2021). *Fairness in machine learning: Lessons from political philosophy*. *ArXiv.org*. <https://doi.org/10.48550/arXiv.1712.03586>
- Bockhorst, J., Conathan, D., & Fung, G. M. (2019). *Probabilistic-logic bots for efficient evaluation of business rules using conversational interfaces*. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(1). <https://doi.org/10.1609/aaai.v33i01.33019422>
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge University Press.
- Brummernhenrich, B., Paulus, C. L., & Jucks, R. (2025). *Applying social cognition to feedback chatbots: Enhancing trustworthiness through politeness*. *British Journal of Educational Technology*. <https://doi.org/10.1111/bjet.13569>
- Diakopoulos, N. (2016). *Accountability in algorithmic decision making*. *Communications of the ACM*, 59(2), 56–62. <https://doi.org/10.1145/2844110>

- Essop, L., Singh, A., & Wing, J. (2023). *Developing a comprehensive evaluation questionnaire for university FAQ administration chatbots*. *2023 Conference on Information Communications Technology and Society (ICTAS)*, 144–151. <https://doi.org/10.1109/ICTAS56421.2023.10082753>
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). *An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations*. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Folger, R., & Cropanzano, R. (2001). *Fairness theory: Justice as accountability*. In J. Greenberg & R. Cropanzano (Eds.), *Advances in Organizational Justice* (pp. 1–55). Stanford University Press.
- Goel, G., Tanwar, P., & Sharma, S. (2022). *UI-UX design using user centred design (UCD) method*. *2022 International Conference on Computer Communication and Informatics (ICCCI)*, 1–8. <https://doi.org/10.1109/ICCCI154379.2022.9740997>
- Kim, D. (2011). *The principle of fairness: Theory, defence, and application*. Warwick.ac.uk. https://wrap.warwick.ac.uk/id/eprint/47260/1/WRAP_THESIS_Kim_2011.pdf
- Langevin, R., Lordon, R. J., Avrahami, T., Cowan, B. R., Hirsch, T., & Hsieh, G. (2021). *Heuristic evaluation of conversational agents*. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–15. <https://doi.org/10.1145/3411764.3445312>
- Liebrecht, C., Sander, L., & van Hooijdonk, C. (2021). *Too informal? How a chatbot's communication style affects brand attitude and quality of interaction*. *Chatbot Research and Design*, 16–31. https://doi.org/10.1007/978-3-030-68288-0_2
- McMahon, C. (2016). *Fairness*. In *Reasonableness and Fairness A Historical Theory* (pp. 21–60). Cambridge University Press. <https://doi.org/10.1017/9781316819340.002>
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). *A survey on bias and fairness in machine learning*. *ACM Computing Surveys*, 54(6), 1–35. <https://doi.org/10.1145/3457607>
- Ononye, C. F. (2020, April 17). *The concept of politeness: First/second order distinction* [Lecture]. Pragmatics Association of Nigeria (PrAN), Telegram platform.
- Pears, M., Henderson, J., & Konstantinidis, S. T. (2021). *Repurposing case-based learning to a conversational agent for healthcare cybersecurity*. In L. Stoicu-Tivadar et al. (Eds.), *Public SHHealth and Informatics: Proceedings of MIE 2021* (Vol. 281, pp. 1066–1070). IOS Press. <https://doi.org/10.3233/SHTI210348>
- Pessoa, L., Martins, L., Hsu, M., & de Freitas, R. (2024). *ZoAM GameBot: A journey to the lost computational world in the Amazonia*. *ACM Journal on Computing and Cultural Heritage*, 17(3), Article 48. <https://doi.org/10.1145/3657303>
- Shahrokhi, M. & Bidabadi, F. (2013). *An overview of politeness theories: Current status, future orientations*. *American Journal of Linguistics*, 2(2), 17–27. https://www.researchgate.net/publication/267948289_An_Overview_of_Politeness_Theories_Current_Status_Future_Orientations
- Song, M., Zhang, H., Xing, X., & Duan, Y. (2023). *Appreciation vs. apology: Research on the influence mechanism of chatbot service recovery based on politeness theory*. *Journal of Retailing and Consumer Services*, 73, 103323. <https://doi.org/10.1016/j.jretconser.2023.103323>
- Spencer-Oatey, H. (2000). *Culturally speaking: Managing rapport through talk across cultures*. London: Continuum.
- Sugisaki, K., & Bleiker, A. (2020). *Usability guidelines and evaluation criteria for conversational user interfaces—A heuristic and linguistic approach*. *MUC 2020: Proceedings of Mensch und Computer 2020*, 309–319. <https://doi.org/10.1145/3404983.3405505>
- Xue, J., Wang, Y.-C., Wei, C., Liu, X., Woo, J., & Kuo, C.-C. J. (2024). *Bias and fairness in chatbots: An overview*. *APSIPA Transactions on Signal and Information Processing*, 13(2). <https://doi.org/10.1561/116.00000064>

FUSION 2025

Doctoral Consortium

Extended Abstract

Extending Interaction Design Dimensions for Augmented Reality Supporting Children with Dyslexia

Emy Hazlinda Mohammad Ridzwan, Husniza Husni
School of Computing, Universiti Utara Malaysia

emy.hazlinda.uum@gmail.com, emyhazlinda_m@ahsgs.uum.edu.my, husniza@uum.edu.my

Introduction and Motivation

Dyslexia is one of the most common learning difficulties among school-aged children, affecting their ability to read, spell, and process information effectively. These challenges often stem from unique cognitive, perceptual, and behavioural characteristics such as phonological deficits, visual confusion, and limitations in working memory (Subramaniam & Kunasegran, 2022). As awareness of dyslexia increases worldwide, there is a growing need for innovative educational technologies that can support more inclusive and effective learning strategies for children with dyslexia.

Augmented reality (AR) has emerged as a promising tool, offering interactive and multisensory experiences that can enhance learner engagement, motivation, and comprehension. By integrating visual, auditory, and interactive elements, AR has the potential to provide more accessible learning pathways for children with dyslexia (Hussain et al., 2023; Lozano-Álvarez et al., 2023). However, despite its potential, most AR applications are designed for mainstream learners. They often prioritise visual aesthetics or basic textual features, while overlooking the distinct cognitive and perceptual ways children with dyslexia perceive and interact with information (Hussain et al., 2023).

These limitations highlight the need to move beyond evaluating AR solely for learning outcomes, instead focusing on the actual interaction experiences of children with dyslexia. Understanding how they engage with AR, where they struggle, where they succeed, and how interaction design (IxD) shapes these experiences is essential for ensuring that AR is not only engaging but also inclusive and effective. This issue highlights the significance of considering IxD dimensions proposed by Silver (2007) and Smith (2004) in this research. Therefore, this research seeks to construct an IxD framework that integrates both established and empirically derived dimensions, aligning AR more closely with the real learning needs of children with dyslexia.

Research Problem

Although AR offers interactive multisensory learning experiences, it is still rarely utilised as an assistive technology in teaching and learning for children with dyslexia. Existing AR applications are typically designed for mainstream learners, emphasising visual appeal or basic textual elements without fully considering the unique cognitive, perceptual, and behavioural characteristics of children with dyslexia. As a result of this suboptimal design, children with dyslexia often encounter a range of interaction difficulties that reduce focus, information processing, and learning effectiveness (Radu & MacIntyre, 2012). Therefore, an urgent need to investigate these interaction challenges and to develop an IxD framework that is more inclusive and aligned with the learning realities of children with dyslexia. This problem arises due to several gaps in existing studies and design practices, which fail to fully address the interaction needs of children with dyslexia in AR contexts.

First, interaction difficulties are underexplored. Many existing AR applications overlook these individual needs due to a lack of in-depth knowledge about the characteristics and learning needs

of children with dyslexia in the context of AR, leading to a mismatch between the learning tool and their unique requirements (Fei et al., 2022). The focus remains largely on textual and visual aspects such as font types and colours. As a result, children with dyslexia frequently encounter confusion with symbols, errors in distinguishing characters, challenges in spatial orientation, and difficulties managing time in tasks, collectively increasing cognitive load and diminishing intended learning benefits (Fei et al., 2022).

Next, although multisensory learning approaches are recognised as effective for children with dyslexia, as they engage multiple senses simultaneously, most AR applications focus mainly on visual and auditory elements (Rello & Baeza-Yates, 2013). The lack of integration of kinaesthetic and tactile inputs prevents AR from fully aligning with the diverse sensory needs of children with dyslexia. The lack of integration of kinaesthetic and tactile modalities prevents AR from fully aligning with the diverse sensory needs of children with dyslexia (Klonari & Passadelli, 2019), thereby limiting the benefits of multisensory engagement.

Moreover, from the literature, most AR are developed with a focus on pedagogical outcome (Bhatti et al., 2020). Such studies typically emphasise outcomes like reading fluency, spelling accuracy, or academic achievement, while overlooking how learners, particularly children with dyslexia, actually interact with and experience the AR environment. This narrow emphasis limits understanding of usability, engagement and interaction flow, which are equally critical for ensuring that AR functions as an effective assistive technology.

A review of studies further reinforces these issues. Fei et al. (2022) found that children with dyslexia often experience confusion and distraction when interacting with AR text and symbols due to overlapping visual stimuli and insufficient spatial guidance. Similarly, Rello and Baeza-Yates (2013) reported that inappropriate font styles and visual overload in digital interfaces can increase cognitive load and reduce reading fluency among dyslexic users. Klonari and Passadelli (2019) highlighted persistent spatial and temporal orientation difficulties, indicating that AR systems rarely account for kinaesthetic engagement. In addition, Bhatti et al. (2020) observed that many AR-based interventions prioritise academic outcomes such as spelling accuracy over user engagement and emotional experience. Collectively, these findings demonstrate that existing AR interaction designs fail to meet the diverse cognitive, perceptual, and emotional needs of children with dyslexia.

These synthesised results from previous literature provide empirical support for the research problem, validating the need for a comprehensive IxD framework that integrates both established and emergent dimensions. Such a framework is necessary to guide the design of inclusive AR learning environments that align with the real interaction needs of children with dyslexia.

Research Objectives

In coherence with the above discussion, the research questions (RQ) are as follows:

RQ1: What are the interaction difficulties that children with dyslexia experience in relation to the IxD dimensions while interacting with AR for learning?

RQ2: How can the identified interaction difficulties be incorporated into a framework tailored for children with dyslexia in AR?

RQ3: To what extent does the proposed framework align with the interaction needs of children with dyslexia?

To achieve this aims, the following research objectives (RO) are outlined:

RO1: To identify the interaction difficulties that children with dyslexia experience in relation to the IxD dimensions while interacting with AR.

RO2: To construct an IxD framework by incorporating suitable IxD dimensions tailored for children with dyslexia in AR.

RO3: To evaluate that the proposed framework is aligned with and relevant to the interaction needs of children with dyslexia.

The RO1 is essential to uncover real interaction barriers faced by children with dyslexia when engaging with AR applications. Understanding these authentic experiences ensures that the framework is grounded in actual learner behaviour rather than assumptions, providing an empirical foundation for inclusive design. Meanwhile, RO2 bridges the gap between theoretical IxD and the practical needs of children with dyslexia. By integrating empirical findings from RO1 with established dimensions, the constructed framework provides a context-specific design guide that reflects both cognitive and multisensory aspects of interaction. RO3 ensures the academic and practical evaluation of the framework through expert review. Evaluating its alignment with learners' interaction needs verifies that the framework is both theoretically coherent and applicable for real-world AR learning environments.

Research Methods

The research design of this research is structured into two sequential phases that align the research objectives with the methods and expected outcomes, as illustrated in Fig.1. This multi-phase design ensures both theoretical coherence and empirical grounding throughout the research process. In Phase 1, a thematic literature review and proof-of-concept prototype were conducted to assess the significance of IxD in AR for children with dyslexia (Ghazali et al., 2023). This preliminary stage establishes the theoretical foundation of the study. It identifies critical gaps in current AR applications, particularly in relation to the unique needs of children with dyslexia. The outcome of this phase justifies the necessity for an in-depth exploration of interaction difficulties and the integration of IxD dimensions into assistive AR design.

In Phase 2, to achieve RO1, observations are conducted to collect data on the interaction difficulties experienced by children with dyslexia when using AR, followed by thematic analysis to generate insights. Observation is chosen because it records genuine learning behaviours and multisensory responses that interviews may overlook (Hassan et al., 2022). Thematic analysis then organises these observations into meaningful patterns (Maguire & Delahunt, 2017), leading to the synthesis of an IxD framework that integrates both established and empirically derived dimensions. While achieving RO2, it focuses on constructing the IxD framework by synthesising findings from RO1 with established IxD dimensions, producing a framework that incorporates suitable design features for AR tailored to the needs of children with dyslexia. Finally, to achieve RO3, the proposed framework is evaluated through expert review. This technique allows experts in education, technology, and dyslexia to assess its alignment, relevance, and practicality (Ssemugabi & Villiers, 2007). Their feedback ensures that the framework is both relevant and applicable. The deliverables of each phase provide a progressive outcome: identifying real

interaction difficulties, producing a tailored IxD framework, and validating its alignment with the learning needs of children with dyslexia.

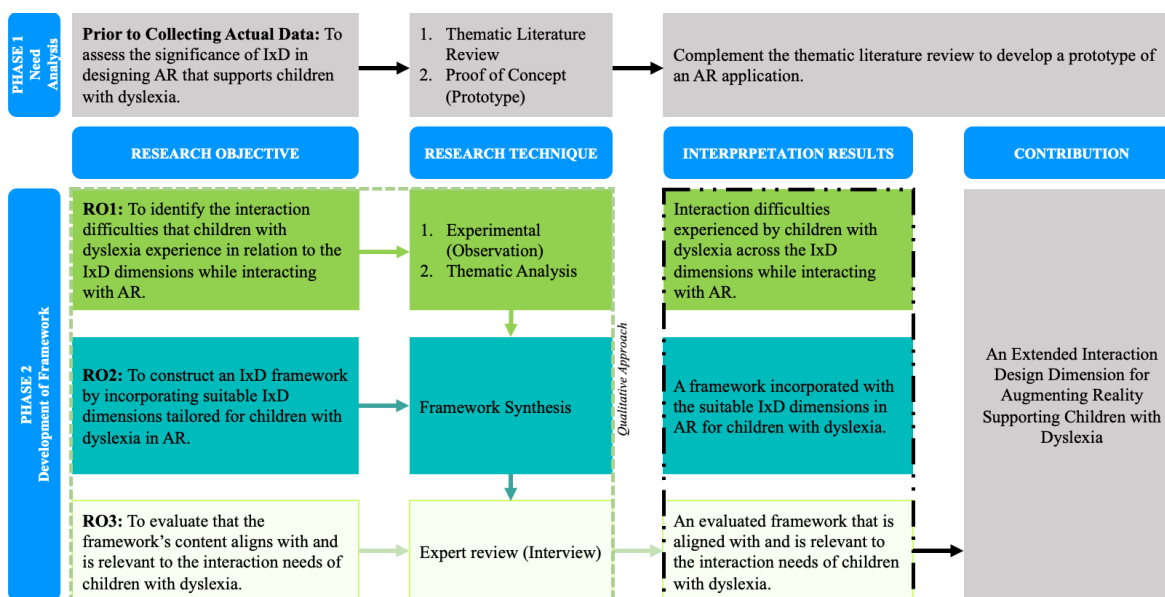


Fig. 1. Research design framework.

Proposed Solution

This research proposes the development of an IxD framework specifically tailored for AR applications that support children with dyslexia. The framework will be constructed by combining established IxD dimensions, as outlined by Silver (Silver, 2007), such as words, visual representations, physical object or space, time, and behaviour, with new dimensions derived inductively from empirical findings, ensuring that both theoretical principles and real-world interaction challenges are addressed. By aligning those design elements and potentially dimensions with the perceptual characteristics of children with dyslexia, the framework will provide structured guidance for design. This solution aims to enhance inclusivity and learning effectiveness, making AR a more reliable assistive technology for children with dyslexia.

Expected Contributions

This research aims to make empirical, theoretical, and practical contributions to the fields of HCI and educational technology. Empirically, it will provide evidence on the specific interaction challenges faced by children with dyslexia in AR environments. Theoretically, it will produce and validate an IxD framework that integrates both established and newly derived dimensions, extending current understanding of IxD in the context of special education. Practically, this research will offer concrete design recommendations for developers, educators, and researchers to create inclusive AR-based educational tools that are responsive to the unique needs of children with dyslexia.

Work to Date

At present, preliminary stages of the research have been completed, including the review of relevant literature, formulation of research questions, and initial development of research instruments. This includes the process of data collection through classroom observations that focus on how children with dyslexia engage with AR tools during learning tasks, and is refined through thematic analysis of empirical findings. The researcher is now developing the draft of the IxD framework, which will later be evaluated through expert review.

Open Issues for Discussion

A key issue for further discussion is whether the dimension of emotion should be formally integrated into the proposed framework. Observations suggest that emotional responses such as anxiety, frustration, or motivation play a critical role in shaping the learning experience of children with dyslexia when interacting with AR. The challenge lies in determining how emotional engagement can be systematically modelled and incorporated without overcomplicating the framework. Additionally, the discussion should address strategies for balancing multisensory integration, ensuring that AR designs are stimulating but not overwhelming for these children.

References

- Bhatti, Z., Shabir, N., Zainab, Z., Ali, L., & Jawaid, A. (2020). Augmented Reality Based Learning System for Children. *2020 3rd International Conference on Computing, Mathematics and Engineering Technologies (ICoMET)*.
- Fei, D., Gao, Z., Yuan, L., & Wen, Z. A. (2022). CollectiAR: Computer Vision-Based Word Hunt for Children with Dyslexia. *CHI PLAY 2022 - Extended Abstracts of the 2022 Annual Symposium on Computer-Human Interaction in Play*, 171–176. <https://doi.org/10.1145/3505270.3558318>
- Ghazali, M., Artha, M. C. S., & Othman, N. Z. S. (2023). A Study on User Interface Design in Learning Arabic Alphabets for Dyslexic Children. *ACM International Conference Proceeding Series*, 76–83. <https://doi.org/10.1145/3604571.3604584>
- Hussain, H., Faisal, N., Saleem, M., Habib, M. A., & Ahmad, M. (2023). Development of an Augmented Reality based Serious Game for the Training of Children with Dyslexia and Dysgraphia. *Lecture Notes in Education Psychology and Public Media (LNEP)*, Vol. 3, 194–200.
- Klonari, A., & Passadelli, A. S. (2019). Differences between dyslexic and non-dyslexic students in the performance of spatial and geographical thinking. *Review of International Geographical Education Online*, 9(2), 284–303. <https://doi.org/10.33403/rigeo.510360>
- Lozano-Álvarez, M., Rodríguez-Cano, S., Delgado-Benito, V., & Mercado-Val, E. (2023). A Systematic Review of Literature on Emerging Technologies and Specific Learning Difficulties. In *Education Sciences* (Vol. 13, Issue 3). MDPI. <https://doi.org/10.3390/educsci13030298>
- Maguire, M., & Delahunt, B. (2017). *Doing a Thematic Analysis: A Practical, Step-by-Step Guide for Learning and Teaching Scholars*. * (Issue 3). <http://ojs.aishe.org/index.php/aishe-j/article/view/335>
- Radu, I., & MacIntyre, B. (2012, November). Using children's developmental psychology to guide augmented-reality design and usability. *IEEE International Symposium on Mixed and Augmented Reality 2012 Science and Technology Proceedings*.

- Rello, L., & Baeza-Yates, R. (2013). Good Fonts for Dyslexia. *ASSETS 2013: Conference Program of the 15th ACM SIGACCESS International Conference on Computers and Accessibility: Bellevue, Washington, USA*, 343.
- Silver, K. (2007). *What Puts the Design in Interaction Design*. <https://www.uxmatters.com/mt/archives/2007/07/>
- Smith, G. C. (2004). *Foreword What Is Interaction Design?*
- Ssemugabi, S., & Villiers, R. de. (2007). A Comparative Study of Two Usability Evaluation Methods Using a Web-Based E-Learning Application. *SAICSIT 2007 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists*, 132–142.
- Subramaniam, V., & Kunasegran, K. (2022). Reading Skills in Malay Language Literacy of Dyslexic Students. *Jurnal Bahasa*, 22(2), 329–346. [https://doi.org/10.37052/jb22\(2\)no7](https://doi.org/10.37052/jb22(2)no7)

Social Media Platforms as Learning Environments: A Pilot Test of Discord and TikTok Live

Nurul Asmaa Abd Razak, Noris Mohd Norowi
Universiti Putra Malaysia, Malaysia

ABSTRACT

Social media integration in the classroom has opened up new avenues for instruction and learning, especially for students who are digital natives. Discord's and TikTok Live's suitability as online learning platforms was investigated in this pilot study. A total of 28 undergraduate students participated in learning sessions conducted on both platforms, followed by a questionnaire and semi-structured interviews. The analysis focused on four dimensions of online learning: engagement and interaction, learning and understanding, collaboration and communication, and accessibility and usability. Results indicate a clear preference for Discord, with 85.7% of students selecting it as the most effective platform, compared to only 10.7% for TikTok Live. According to qualitative feedback, TikTok Live was viewed as limited and engaging, better suited for asynchronous or microlearning use, while Discord excelled at promoting collaboration, content organization, and interactivity. The results highlight how crucial it is to match platform affordances to students' digital practices and pedagogical objectives. Future studies will build on this work by contrasting Microsoft Teams with Discord and investigating the effects of long-form (YouTube) versus short-form (TikTok) video learning.

Keywords: Discord, TikTok Live, Online Learning, Social Media in Education, Student Engagement, Collaboration and Communication, Microlearning, Digital Pedagogy

INTRODUCTION

As digital technologies have grown rapidly, they have changed not only how people interact and communicate but also how educators create and present educational experiences. Social media platforms, which were initially created for networking and entertainment purposes, have become more and more integrated into education as students look for flexible, interesting, and cooperative learning environments outside of the conventional classroom (Selwyn, 2021; Kumar & Nanda, 2024). Finding platforms that support academic goals and align with students' digital habits has become even more crucial as a result of the global shift toward online learning, which has been accelerated by the COVID-19 pandemic (Bao, 2020; Dhawan, 2020).

According to Adedoyin and Soykan (2023), traditional online learning platforms like Zoom, Microsoft Teams, and Google Meet have demonstrated their dependability in synchronous instruction, evaluation, and communication. However, their structured, institution-oriented designs may not fully align with the participatory culture and dynamic interaction patterns that characterize students' everyday digital practices (Mannell & Smith, 2022). Younger generations, on the other hand, have taken to community-driven platforms like Discord (Arifianto & Izzudin, 2021) and live-streaming services like TikTok Live (Opara et al., 2025), which frequently turn into unofficial venues for peer support, collaboration, and knowledge sharing. These platforms'

emphasis on accessibility, multimodal communication, and real-time interaction make them especially appealing and suggest that they could be adapted for use in educational settings (Ahmad, 2024).

Despite this potential, there remains a lack of empirical evidence on whether such platforms can effectively support structured learning, particularly when compared with established tools (Dron & Anderson, 2014; Archambault & Rice, 2022). Due to worries about usability, information flow, and appropriateness for academic settings, many educators are still reluctant to use social media platforms for academic purposes (Al-Hail et al., 2024). As such, pilot studies are critical in identifying which platforms can meaningfully complement or enhance existing educational technologies. This study examines the educational suitability of TikTok Live and Discord, concentrating on four aspects that are essential to online learning: communication and collaboration, learning and comprehension, accessibility and usability, and engagement and interaction.

BACKGROUND

Social Media in Education

In recent years, social media has become increasingly integrated into education, offering both formal and informal learning opportunities (Kumar & Nanda, 2024). Beyond their use as communication and entertainment platforms, platforms such as YouTube, Facebook, and Instagram have been used for peer learning, academic engagement, and knowledge sharing (Kazmi et al., 2024; Ovhal, 2025; Joshi & Kumar, 2025). The use of digital platforms for education increased during the COVID-19 pandemic, and social media emerged as an additional avenue for communication and information sharing (Dwivedi et al., 2020). Research shows that social media, especially when combined with official learning management systems, improves student motivation, engagement, and collaborative learning (Al-Rahmi et al., 2021). Still, there are problems with the shared information's dependability, privacy, and distraction.

Discord as a Learning Tool

Originally launched as a platform for online gaming communities, Discord has evolved into a versatile communication tool with significant potential for education (AlGhamdi, 2025). Collaborative learning and peer interaction are well-aligned with its features, which include voice and video conferencing, persistent chat channels, and community-building servers (Reitman et al., 2021; Ling, 2022). According to recent studies, Discord helps students feel like they belong and are part of a community, which makes it particularly useful for group projects, conversations, and peer-to-peer mentoring (AlGhamdi, 2025; Kazmi et al., 2024). Additionally, compared to institutional platforms, its informal setting lowers barriers to participation, allowing students to participate more freely. During the pandemic, educators reported using Discord to support flipped classrooms, office hours, and collaborative assignments (Hew et al., 2020). These results imply that Discord is a useful learning management system when modified for educational purposes in addition to being a social tool.

TikTok in Educational Contexts

TikTok, a short-form video platform, has rapidly grown into one of the most popular apps among younger audiences. Researchers and educators have taken notice of its potential for microlearning, which involves imparting knowledge through brief, captivating videos (Khlaif & Salha, 2021; Conde-Caballero et al., 2024). Recent studies highlight that TikTok supports quick information sharing, creativity, and learner engagement, especially in informal or supplementary learning contexts (Kaye et al., 2021; Nguyen & Diederich, 2023; Al-Azawei & Alowayr, 2025). TikTok has been used by educators to provide language learning materials (Thorvik, 2023; Sarkila et al., 2024; Alharthy, 2025), scientific explanations (Shaafi et al., 2023; Abualrob, 2025), and inspirational content (Yélamos-Guerra et al., 2022), which capitalize on the platform's ability to attract attention and enhance recall. However, there are drawbacks, such as the platform's culture of entertainment, the possibility of false information, and its limited ability to support in-depth, structured learning (Horgan et al., 2024). Despite these limitations, TikTok continues to show promise as an additional tool for promoting engagement and increasing access to educational resources.

METHODOLOGY

This study adopted a mixed-method pilot research design to evaluate the suitability of two digital platforms—Discord and TikTok Live—for educational use. A pilot study was selected because, prior to a full-scale study, it allows researchers to test tools, detect possible problems, and improve the research framework (Teresi et al., 2022; Muasya & Mulwa, 2023). According to Takona (2024), pilot testing is especially useful in digital education research because platform affordances and student experiences can differ greatly depending on the context.

Participants

A total of 28 undergraduate computer science students willingly took part in this study. Students from the Computer Science programme were selected for this pilot test in order to guarantee that the findings reflected both technical feasibility and user experience, as well as to set the groundwork for future research that might include a more diverse student body. This ensured that the study was relevant to the focus on digital learning practices.

Procedure

Two separate teaching sessions were conducted, each utilizing one of the platforms—Discord and TikTok Live. To ensure comparability, the same instructional content was delivered across all platforms. After each session, participants completed a questionnaire structured around four key dimensions of online learning, which are engagement & interaction, learning & understanding, collaboration & communication, and accessibility & usability. These categories align with current frameworks for online interaction and learning (Martin et al., 2023; Alam & Mohanty, 2023).

Data Collection

Data were collected through two instruments:

- **Quantitative Survey:** A short questionnaire captured students' perceptions using Likert-scale items (South et al., 2022).
- **Qualitative Interviews:** Semi-structured interviews were conducted immediately after the teaching sessions to obtain richer insights into students' preferences and experiences (Mutanga, 2024).

To strengthen the validity of the results and triangulate findings, this convergent mixed-methods approach was used (Plano Clark, 2017; Ye & Hutson, 2023; Sharma et al., 2023).

Data Analysis

Descriptive analysis of quantitative data was done, with an emphasis on preference rankings and frequency distributions (Ghanad, 2023). Then, thematic analysis of the qualitative data revealed recurrent trends in the ways that students encountered usability, collaboration, and engagement across platforms (Squires, 2023).

A balanced understanding of the breadth (survey data) and depth (interview data) of student experiences was made possible by the combination of descriptive statistics and thematic coding.

RESULT & DISCUSSION

A total of 28 undergraduate students participated in the pilot test, which explored the effectiveness of Discord and TikTok Live as platforms for educational use. After each teaching session, students completed a questionnaire and participated in interviews to reflect on their learning experiences. The analysis focused on four dimensions: (A) Engagement & Interaction, (B) Learning & Understanding, (C) Collaboration & Communication, and (D) Accessibility & Usability.

Platform Experience and Usage

- **Discord:** All 28 participants had prior experience with Discord. As seen in Fig. 2, 16 (57.1%) said they had previously used it for educational purposes, whereas 12 (42.9%) had only used it for social or recreational purposes.
- **TikTok Live:** 5 (17.9%) participants had never used the platform before, while 23 (82.1%) participants reported having prior experience, as shown in Fig. 1. 22 (81.5%) students reported using TikTok Live for educational or informational purposes, as shown in Fig. 2.

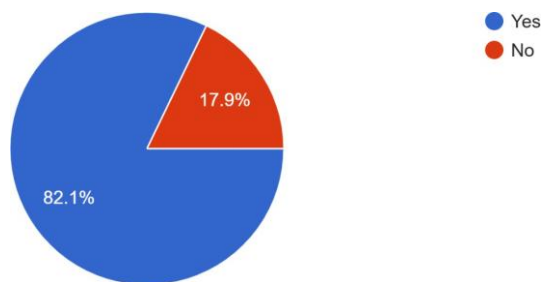


Fig. 1: Chart for the question, 'Do you have experienced using TikTok Live or join any TikTok live?'

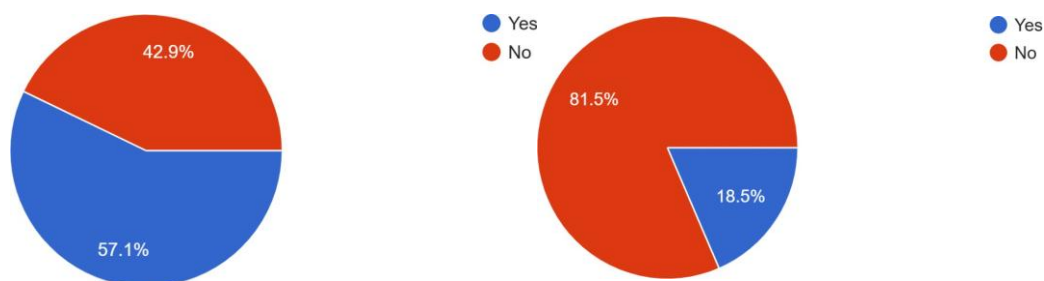


Fig. 2: Chart for the question, [Left]: 'Do you use Discord for learning purposes?', [Right]: 'Do you use TikTok Live for learning purposes?'

Platform Preference

In response to the question of which platform they thought was best for learning, most students chose Discord. Out of the 28 students who responded, 24 (85.7%) chose Discord as their platform of choice, 3 (10.7%) chose TikTok Live, and 1 (3.6%) said they had no preference (Table 1).

Table 1: Platform Preference (n = 28).

Platform	Students	Percentage
Discord	24	85.7%
TikTok Live	3	10.7%
None	1	3.6%

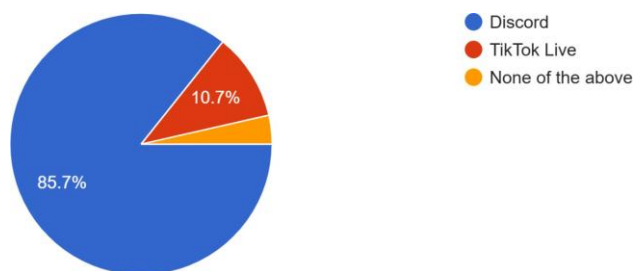


Fig. 3: Chart for the question, 'Between Discord and TikTok Live, which one do you prefer for the learning purposes?'

According to this distribution, Discord was widely thought to be more useful for facilitating organized learning activities. Students' strong preference for Discord was reaffirmed during the interviews, where they highlighted its collaborative affordances, organizational features, and interactive nature. Many participants emphasised how Discord enabled them to share resources within specific channels, revisit previously published content, and communicate in multiple formats (text, voice, and video). These features helped people view Discord as a community-based learning environment as opposed to a one-way teaching aid.

By contrast, only three students selected TikTok Live, recognizing its advantages in terms of immediate engagement but also pointing out important drawbacks. It was stated that TikTok Live was better suited for informal communication or quick learning exercises than for lengthy class periods. Students believed that its value as a primary educational platform was limited by its one-way communication style, lack of presentation features, and high potential for distraction.

These results are consistent with more general research in digital pedagogy that shows students prefer platforms that combine flexibility, interactivity, and persistent content over those that are mainly used for broadcasting or entertainment (Bond et al., 2021; Fields & Kafai, 2021).

Thematic Insights

1. Engagement & Interaction

Discord was frequently cited by students as being more participatory and interesting than TikTok Live. They felt that having access to a variety of communication tools, including voice channels, screen sharing, text chat, and even customized soundboards, strengthened the sense of community during class. On the other hand, TikTok Live was thought to be primarily one-way, with students characterizing it as "just the lecturer talking and students commenting," frequently interrupted by technical difficulties. Additionally, several participants reported that TikTok's entertainment-focused features (pop-ups, tagging, and GIFs) created distractions that hindered learning focus.

While Fields and Kafai (2021) argue that platforms like Discord increase student participation by fostering a community-oriented learning culture, Opara et al. (2025) caution that TikTok's design promotes attention but not sustained academic engagement.

2. Learning & Understanding

Students highlighted how Discord's channels enabled them to review discussions, slides, and instructions asynchronously, which promoted deeper understanding. This was especially helpful when reviewing explanations or getting ready for assignments. TikTok Live, on the other hand, was thought to be more appropriate for brief, superficial learning—"storytelling" or brief demonstrations—than for in-depth subjects. In contrast to live teaching sessions, some students proposed that TikTok might function better when used asynchronously (pre-recorded short videos, like YouTube).

These opinions align with those of Bond et al. (2021), who highlight the significance of retrievable digital content in enhancing knowledge retention, and Pitaloka (2025), who argue that TikTok is more effective for microlearning than for longer academic delivery.

3. Collaboration & Communication

One of the main reasons why students preferred Discord was its collaborative affordances. Participants explained how it was appropriate for group presentations and peer-to-peer learning because it allowed several students to share screens at once, and peers could select whose feed to follow. Additionally, Discord made file sharing and threaded discussions possible, facilitating teamwork outside of live sessions. In comparison, TikTok Live was perceived as broadcast-focused, offering little chance for in-the-moment interaction or feedback beyond brief remarks.

This corroborates the findings of Martin et al. (2020), who highlight the importance of both synchronous and asynchronous collaboration in digital learning environments, and Banson & Hardin (2022), who found that Discord enhances teamwork and student-to-student engagement.

4. Accessibility & Usability

Students were familiar with both platforms, but Discord was thought to be more useful and adaptable for educational settings. Discord was especially useful for structured learning activities on PCs because of its persistent channels and servers that could be customized. Although TikTok Live was easier to use on mobile devices, according to some participants, this benefit was overshadowed by its lack of security, limited functionality, and distractions. Students were worried that outsiders might join TikTok Live sessions, which made them feel intimidated. In contrast, Discord was thought to be safer because only students could access it.

These findings are consistent with Dwivedi et al. (2020), who emphasize the importance of platforms that strike a balance between security and accessibility while pointing out that privacy and distraction are major obstacles when using open social platforms for education.

Comparative Evaluation

The comparative analysis reveals the divergent advantages and disadvantages of TikTok Live and Discord in relation to four important aspects of online learning: accessibility & usability, learning & comprehension, collaboration & communication, and engagement & interaction. Table 2 provides a clearer view.

Table 2: Comparative Strengths and Weaknesses of Discord and TikTok Live.

Dimension	Discord (Strengths)	Discord (Weaknesses)	TikTok Live (Strengths)	TikTok Live (Weaknesses)
Engagement & Interaction	Multimodal, fosters community	May require orientation for new users	Captures attention quickly	One-way, delayed, distracting features
Learning & Understanding	Organized, retrievable content	Can become cluttered without moderation	Effective for microlearning	Ephemeral, lacks depth
<i>Collaboration & Communication</i>	<i>Group projects, peer-to-peer sharing</i>	Dependent on community norms for effectiveness	<i>Quick Q&A possible</i>	<i>No real collaboration tools</i>
<i>Accessibility & Usability</i>	<i>Flexible, customizable, multi-device support</i>	Some functions easier on PC than mobile	<i>Ubiquitous, mobile-friendly</i>	<i>Not suited for structured tasks, privacy concerns</i>

CONCLUSION

This pilot test shows that, among the platforms tested, Discord is the best for education, outperforming TikTok Live in terms of collaboration, engagement, learning support, and usability. Although TikTok Live has the potential to be used for asynchronous content and microlearning, it does not have the organizational and collaborative tools required for formal academic learning.

The study highlights how crucial it is to choose platforms that promote active engagement, knowledge retention, and secure learning environments in addition to grabbing users' attention. By contrasting Discord with Microsoft Teams and investigating the effects of short-form versus long-form video learning, future research will broaden.

REFERENCES

- Abualrob, M. (2025). TIKTOK FOR SCIENCE LEARNING: THE INTERPLAY OF TIKTOK AS AN EDUCATIONAL TOOL, USABILITY, SATISFACTION, SKILLS, AND FUTURE IMPACT.
- Adedoyin, O. B., & Soykan, E. (2023). Covid-19 pandemic and online learning: the challenges and opportunities. *Interactive learning environments*, 31(2), 863-875.
- Ahmad, E. A. (2024). Revolutionizing learning: leveraging social media platforms for empowering open educational resources. *International Journal of e-Learning and Higher Education (IJELHE)*, 19(1), 83-106.
- Alam, A., & Mohanty, A. (2023). Educational technology: Exploring the convergence of technology and pedagogy through mobility, interactivity, AI, and learning tools. *Cogent Engineering*, 10(2), 2283282.
- Al-Azawei, A., & Alowayr, A. (2025). Predicting the Acceptance of Informal Learning Technologies: A Case of the TikTok Application. *Education Sciences*, 15(3), 362.
- AlGhamdi, R. (2025). Bridging learning gaps through Discord: peer-to-peer learning in computer graphics education. *Learning and Teaching in Higher Education: Gulf Perspectives*.
- Al-Hail, M., Zguir, M. F., & Koç, M. (2024). Exploring digital learning opportunities and challenges in higher education institutes: Stakeholder analysis on the use of social media for effective sustainability of learning–teaching–assessment in a university setting in Qatar. *Sustainability*, 16(15), 6413.
- Alharthy, F. (2025). Exploring the Impact of TikTok on Second-Language Vocabulary Acquisition: Benefits, Challenges, and Learner Perceptions. *Journal of Humanities and Social Sciences Studies*, 7(3), 22-31.
- Al-Rahmi, A. M., Shamsuddin, A., Alturki, U., Aldraiweesh, A., Yusof, F. M., Al-Rahmi, W. M., & Aljeraiwi, A. A. (2021). The influence of information system success and technology acceptance model on social media factors in education. *Sustainability*, 13(14), 7770.
- Archambault, L., Leary, H., & Rice, K. (2022). Pillars of online pedagogy: A framework for teaching in online learning environments. *Educational Psychologist*, 57(3), 178-191.
- Arifianto, M. L., & Izzudin, I. F. (2021). Students' Acceptance of Discord as an Alternative Online Learning Media. *International Journal of Emerging Technologies in Learning*, 16(20), 179–195. <https://doi.org/10.3991/ijet.v16i20.22917>
- Banson, J., & Hardin, C. D. (2022, June). Assessing student participation and engagement using discord. In *2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC)* (pp. 1299-1305). IEEE.
- Bao, W. (2020). COVID-19 and online teaching in higher education: A case study of Peking University. *Human Behavior and Emerging Technologies*, 2(2), 113–115. <https://doi.org/10.1002/hbe2.191>
- Bond, M., Bedenlier, S., Marín, V. I., & Händel, M. (2021). Emergency remote teaching in higher education: Mapping the first global online semester (Pre-print).
- Conde-Caballero, D., Castillo-Sarmiento, C. A., Ballesteros-Yáñez, I., Rivero-Jiménez, B., & Mariano-Juárez, L. (2024). Microlearning through TikTok in Higher Education. An evaluation of uses and potentials. *Education and Information Technologies*, 29(2), 2365–2385. <https://doi.org/10.1007/s10639-023-11904-4>
- Dhawan, S. (2020). Online learning: A panacea in the time of COVID-19 crisis. *Journal of Educational Technology Systems*, 49(1), 5–22. <https://doi.org/10.1177/0047239520934018>
- Dron, J., & Anderson, T. (2014). Teaching crowds: Learning and social media. *Athabasca University Press*.
- Dwivedi, Y. K., Hughes, D. L., Coombs, C., Constantiou, I., Duan, Y., Edwards, J. S., ... & Upadhyay, N. (2020). Impact of COVID-19 pandemic on information management research and practice: Transforming education, work and life. *International journal of information management*, 55, 102211.

- Fields, D. A., & Kafai, Y. B. (2021). Youth learning through making and gaming on Discord: Understanding an emerging digital learning ecology. *Learning, Media and Technology*, 46(4), 466–478. <https://doi.org/10.1080/17439884.2021.1900241>
- Ghanad, A. (2023). An overview of quantitative research methods. *International journal of multidisciplinary research and analysis*, 6(08), 3794-3803.
- Hew, K. F., Jia, C., Gonda, D. E., & Bai, S. (2020). Transitioning to the “new normal” of learning in unpredictable times: pedagogical practices and learning performance in fully online flipped classrooms. *International Journal of Educational Technology in Higher Education*, 17(1), 57.
- Horgan, B. L., Shine, J., & Farrelly, T. (2024). The Modern Scroll: Exploring TikTok as a Platform for Heutagogical Learning.
- Joshi, P., & Kumar, D. (2025). Social Media Transforming Higher Education Beyond Likes and Shares. *Telangana Journal of Higher Education*, 1(1), 43-47.
- Kazmi, S., Jan, R., Shaikh, H. L., & Javeed, S. (2024). How Educational Institutions use Social Media Platforms to Engage Students, Promote Online Communities and Build a Sense of Belonging. *Review of Education, Administration & Law*, 7(4), 303-317.
- Kaye, D. B. V., Chen, X., & Zeng, J. (2021). The co-evolution of two Chinese mobile short video apps: Parallel platformization of Douyin and TikTok. *Mobile Media & Communication*, 9(2), 229-253.
- Khlaif, Z. N., & Salha, S. (2021). Using TikTok in Education: A Form of Micro-learning or Nano-learning? *Interdiscip J Virtual Learn Med Sci*, 12(3), 213–218. <https://doi.org/10.30476/ijvlms.2021.90211.1087>
- Kumar, V., & Nanda, P. (2024). Social media as a learning tool: A perspective on formal and informal learning. *International Journal of Educational Reform*, 33(2), 157-182.
- Ling, G. (2022). Learning with your Buddies: an investigation of community based UX design learning on Discord.
- Mannell, K., & Smith, E. T. (2022). Alternative social media and the complexities of a more participatory culture: A view from Scuttlebutt. *Social Media+ Society*, 8(3), 20563051221122448.
- Martin, F., Sun, T., & Westine, C. D. (2020). A systematic review of research on online teaching and learning from 2009 to 2018. *Computers & education*, 159, 104009.
- Martin, F., Bacak, J., Polly, D., & Dymes, L. (2023). A systematic review of research on K12 online teaching and learning: Comparison of research from two decades 2000 to 2019. *Journal of Research on Technology in Education*, 55(2), 190-209.
- Muasya, J. N., & Mulwa, P. M. (2023). Pilot study, a neglected part of qualitative and quantitative research process: Evidence from selected PhD thesis and dissertations. *Higher Education Research*, 8(4), 115-123.
- Mutanga, M. B. (2024). Students' perspectives and experiences in project-based learning: A qualitative study. *Trends in Higher Education*, 3(4), 903-911.
- Nguyen, H., & Diederich, M. (2023). Facilitating knowledge construction in informal learning: A study of TikTok scientific, educational videos. *Computers & Education*, 205, 104896.
- Opara, E., Adalikwu, T. M.-E., & Tolorunleke, C. A. (2025). The Impact of Tiktok's Fast-Paced Content on Attention Span of Students. <https://doi.org/10.20944/preprints202501.0269.v1>
- Ovhal, D. R. (2025). EVOLUTION OF SOCIAL MEDIA: AS INNOVATIVE RESEARCH AND EDUCATIONAL TOOL. *The Infinite Path: Research Across Disciplines*, 136.
- Pitaloka, N. (2025). The Impact of Video-Based Microlearning on Student Learning Outcomes and Motivation. *Journal of Education Innovation and Curriculum Development*, 3(2), 72-79.
- Plano Clark, V. L. (2017). Mixed methods research. *The Journal of Positive Psychology*, 12(3), 305-306.
- Reitman, J. G., Anderson, C. G., & Steinkuehler, C. (2021). Discord community challenges, tools, and strategies. Irvine, CA: Connected Learning Lab.

- Sarkila, D., Islam, M. H., & Fatmawati, R. (2024). Students Perception On The Use Of Tiktok For Learning Speaking English. *Fonologi: Jurnal Ilmuan Bahasa Dan Sastra Inggris*, 2(2), 01-14.
- Selwyn, N. (2021). Education and technology: Key issues and debates. *Bloomsbury Publishing*.
- Sharma, L. R., Bidari, S., Bidari, D., Neupane, S., & Sapkota, R. (2023). Exploring the mixed methods research design: types, purposes, strengths, challenges, and criticisms. *Global Academic Journal of Linguistics and Literature*, 5(1), 3-12.
- Shaafi, N. F., Mohd Yusof, M. M., Mohammad Khalipah, N. N., & Mohd Hanif, N. (2023). Investigating TikTok as a learning tool for learning chemistry: a study among secondary school students in Malaysia.
- South, L., Saffo, D., Vitek, O., Dunne, C., & Borkin, M. A. (2022, June). Effective use of Likert scales in visualization evaluations: A systematic review. *In Computer Graphics Forum* (Vol. 41, No. 3, pp. 43-55).
- Squires, V. (2023). Thematic analysis. In *Varieties of qualitative research methods: Selected contextual perspectives* (pp. 463-468). *Cham: Springer International Publishing*.
- Takona, J. P. (2024). Research design: qualitative, quantitative, and mixed methods approaches. *Quality & Quantity*, 58(1), 1011-1013.
- Teresi, J. A., Yu, X., Stewart, A. L., & Hays, R. D. (2022). Guidelines for designing and evaluating feasibility pilot studies. *Medical care*, 60(1), 95-103.
- Thorvik, E. G. (2023). Is TikTok a New Way to Support Language Learning (Bachelor's thesis, NTNU).
- Yélamos-Guerra, M. S., García-Gámez, M., & Moreno-Ortiz, A. J. (2022). The use of Tik Tok in higher education as a motivating source for students. *Porta Linguarum Revista Interuniversitaria de Didáctica de las Lenguas Extranjeras*, (38), 83-98.
- Ye, H., & Hutson, B. L. (2023). Using Mixed Methods. In *Scholarly Inquiry in Academic Advising* (pp. 154-175). Routledge.

Collaborative Emotion-Recognition Brain-Computer Interface for User-Centric UX Evaluation

Ting Ying Lai, Masitah Ghazali
Faculty of Computing, Universiti Teknologi Malaysia, Johor Bahru, Malaysia
tingying@graduate.utm.my

ABSTRACT

Despite significant advancements in deep learning-driven Artificial Intelligence (AI), a critical gap exists in fields related to Human-Computer Interface (HCI), such as application interface design and objective evaluation of affective user experience (UX). Conventional UX practices often rely on subjective feedback and observation, failing to capture users' genuine emotional states. This study introduces an electroencephalography (EEG)-based collaborative emotion-recognition brain-computer interface (cER-BCI) model to detect users' real-time emotions, acting as a supportive collaborator that allows UX practitioners to prompt timely interventions or alternative designs. This research aims to establish an efficient and accurate collaborative emotion recognition framework and evaluate its effectiveness in promoting emotional awareness within UX simulations. The proposed methodology involves training the ER model using CCNN on the DEAP dataset, developing a visualization dashboard, and validating its effectiveness through expert interviews and UX simulations. This work seeks to establish a more empathetic and data-informed standard for human-centered UX practices.

Introduction

Motivation

Conventional user experience (UX) evaluation methods rely on behavioral observation and post-hoc self-reporting, which struggle to capture users' genuine affective states in real time. EEG-based emotion recognition allows for a more streamlined real-time observation of a user's feelings while interacting with a given application. The aim is not to automate design-related decisions but to support human evaluators with timely, interpretable affective information, enhancing understanding and empathy in UX analysis, which eventually leads to better, human-centered interface design and experiences.

Problem statement, research questions, and objectives

The user's emotional state is one of the direct factors impacting user satisfaction and product success [0]. Nevertheless, conventional UX evaluation methods, such as post-hoc self-reporting and subjective behavioral observation, may suffer from different challenges and biases. For instance, recall bias, observer bias, language barriers, and an inability to capture users' authentic emotional responses in real-time [0, 0].

In addition, many current practices tend to focus on the usability and functionality of applications, for example, ease of use and easy to learn, overlooking the importance of exploring users' true emotions [0, 0]. This can lead to usable products with low user satisfaction and potentially emotional disconnection, failing to foster engagement in the long term. As digital devices and interfaces increasingly influence daily life, it is crucial to create UX contexts that are more emotion-aware and human-centered.

Furthermore, despite the advancement of deep learning in classifying emotions using physiological signals, a significant gap remains in collaborative applications within the Human-

Computer Interface (HCI) field, specifically for UX evaluation. There is a lack of a collaborative framework that translates real-time EEG data into interventions (e.g., feedback) to establish an efficient and user-centric UX environment.

This study defines the following research questions (RQs) and objectives (ROs) (see **Table 1**).

Table 1. Research questions and objectives.

Research question	Research objective
How to incorporate real-time interventions to facilitate emotional awareness and collaboration between users and UX practitioners during evaluations?	To design a cER-BCI model for affective communication and intervention between users and UX practitioners during evaluations.
How to evaluate the accuracy of the cER-BCI model and present the detected emotions for UX contexts?	To evaluate the accuracy of the cER-BCI model and visualize the emotions detected for UX contexts.
How to assess the effectiveness of the cER-BCI model in supporting UX evaluators in making informed, ethical, and human-centered interventions.	To assess the effectiveness of the cER-BCI model in supporting UX evaluators in making informed, ethical, and human-centered interventions.

Background and related work

Current practices in UX lack affective computing, limiting the technology to post-session analysis and passive monitoring. The approaches are often passive, where users are usually given different tasks while using the applications, while UX evaluators observe the facial expressions and body movements of users, which creates passive interactions during the sessions. For post-session analysis, surveys, questionnaires, and interviews are usually conducted to gather feedback from users or consumers for evaluating various aspects of UX, including usability, functionality, accessibility, and so on [0].

With the invention of artificial intelligence, deep learning models have been developed for a variety of applications, utilizing different algorithms with varying degrees of success in recognizing patterns. This advancement allows further interpretation of electrical signals generated by the brain, such as EEG signals, that reflect human emotions. Along with this, ER and BCI systems have been gaining increasing attention from researchers over the years for their enhanced portability and applicability. However, most of the previous studies focus on accuracy improvement, leaving their implementations in UX practices and collaborative contexts underexplored [0, 0, 0].

Research methodology

This study focuses on providing a more objective solution to the UX practices other than the conventional approaches (e.g., questionnaires and observations) by introducing the EEG-based collaborative emotion-recognition brain-computer interface (cER-BCI) model as a collaborator supporting the UX practitioners, rather than a decision-maker, in establishing a more empathetic and emotion-aware environment, which helps improve overall user experience. The ER is modelled to detect real-time emotions using EEG collected through a consumer-grade brain-computer interface (BCI) headset while users interact with different interfaces during the UX

evaluation sessions. Meanwhile, the evaluators are able to observe both live and emotional results from the ER model to understand the users better and in-depth, which aids in making decisions on whether they should provide immediate intervention, such as feedback or alternative designs. The goal of this research is to streamline the UX practice process by incorporating physiological signals—EEG signals, which can capture hidden and true affective states—into conventional evaluation methods, including behavioral cues and expert interpretations.

Phase 1: Collaborative model design (to achieve RQ1 and RO1)

This phase included a literature review that helped to identify the gap between the deep learning-driven Artificial Intelligence (AI) and user experience (UX) practices, specifically in prompting immediate interventions and establishing collaborative interactions between users and evaluators. This literature review also includes exploring the implementation of deep learning and emotion recognition utilizing different physiological signals, such as EEG and facial muscle movements. Furthermore, a UX evaluation observation session was held by the industry collaborator, which provided informative insights on how UX evaluations are conducted in real-world environments. Apart from proving the application of conventional UX approaches, the findings from both the literature review and the UX practice observation contribute to the humanized intelligence-driven design of the cER-BCI model, which supports UX evaluators in emotional awareness and decision-making (see Fig. 2).

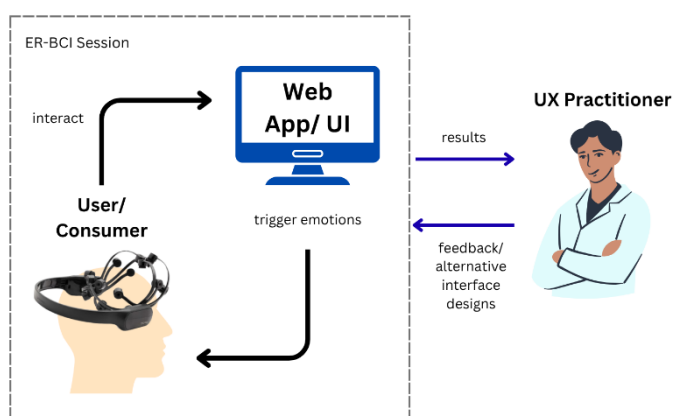


Fig. 2. cER-BCI model with real-time intervention

Phase 2: Emotion recognition training, testing, and visualization (to achieve RQ2 and RO2)

DEAP (Database for Emotion Analysis using Physiological Signals), one of the most popular emotion datasets, is utilized to train and test the ER-BCI model using CCNN from TorchEEG. The emotional data recorded in DEAP are known as valence (positive or negative emotion) and arousal (the intensity of the emotion), which can represent different emotions targeted in this study: happy, sad or bored, neutral, and frustrated. The emotion recognition model evaluation employs confusion-matrix analysis and performance metrics, such as accuracy and F1-score, to ensure its reliability and applicability. Additional validation will be conducted in the UX simulation phase to compare emotions detected with self-reported ratings from users to ensure ecological relevance within UX contexts.

Moreover, a real-time dashboard will be developed to visualize emotional states detected in real time. Instead of showing emotions recognized in the forms of valence and arousal, direct emotion

representation will be displayed for easy and quick simultaneous analysis and interpretation, especially when evaluators also have to pay attention to users' behavioral cues. For example, high valence and high arousal will be replaced by the word "happy" with an emoji.

Phase 3: User experience (UX) simulations (to achieve RQ3 and RO3)

To investigate the effectiveness of the proposed cER-BCI model, experiments that simulate UX evaluation sessions will be carried out. In each experiment, while wearing the BCI headset, the user will be required to complete different tasks assigned on a web application that is built to trigger various emotions intentionally. Real-time emotional states will be collected, detected, and visualized on the dashboard to be interpreted by UX evaluators. Then, evaluators can decide to prompt immediate suitable interventions for users. After the simulations, the emotions detected will be validated with the users to assess the accuracy of the cER-BCI model. Interviews will also be conducted to evaluate the impact of emotional data insights and real-time interventions on overall user experience.

Expected Contribution

It is expected that this proposed cER-BCI model will act as a collaborator, a supporting role, with the UX practitioners during the evaluation process. The practitioners can combine both emotional insights from the model and their own expert judgement to make informed and ethical decisions during the UX sessions, providing real-time interventions (e.g., feedback and alternative interface designs), which establishes a more engaging and user-centric environment between the users and UX practitioners, pushing UX evaluation methods towards a humanized intelligence-driven design.

Issues for discussion

Some challenges in this research include:

Integrating cER-BCI with existing UX practices: how can the collaborative cER-BCI system be best integrated into the established UX practices without causing excessive cognitive load to evaluators?

Emotion recognition challenge: given that the BCI model is trained on controlled laboratory datasets (DEAP), what advanced data processing or fine-tuning techniques should be prioritized to maintain classification accuracy and reliability in a real-world and real-time setting with the presence of noise, artifacts, and variant physiological signals?

Balancing automation and human judgment: what are the appropriate interventions that ensure the cER-BCI remains a supporting role rather than replacing the UX practitioner's critical thinking and judgment?

References

- Marques, L., Matsubara, P. G., Nakamura, W. T., Ferreira, B. M., Wiese, I. S., Gadelha, B. F., Zaina, L. M., Redmiles, D., & Conte, T. U. (2021). Understanding UX Better: A New Technique to Go beyond Emotion Assessment. *Sensors*, 21(21), 7183. <https://doi.org/10.3390/s21217183>
- Snow, P., Ruiz-Segura, A., Léger, P., Sénécal, S., Coursaris, C., Pourchon, R., Cosby, S., & Beauchesne, A. (2025). Breaking the bias: Integrating Physiological and Self-Reported data to improve UX researchers' accuracy and empathy. *Computers in Human Behavior Reports*, 19, 100723. <https://doi.org/10.1016/j.chbr.2025.100723>

- Maia, C. L. B., & Furtado, E. S. (2016). A systematic review about user experience evaluation. *Lecture Notes in Computer Science*, 445–455. https://doi.org/10.1007/978-3-319-40409-7_42
- Ntoa, S. (2024). Usability and User experience evaluation in Intelligent Environments: A review and reappraisal. *International Journal of Human-Computer Interaction*, 1–30. <https://doi.org/10.1080/10447318.2024.2394724>
- Ghani, M. S. a. A., & Shamsuddin, S. N. W. (2020). DEFINITIONS AND CONCEPTS OF USER EXPERIENCE (UX): A LITERATURE REVIEW. *Deleted Journal*, 8(1), 130–143. <https://doi.org/10.47252/teniat.v8i1.292>
- Baek, H. J., Chang, M. H., Heo, J., & Park, K. S. (2019). Enhancing the usability of Brain-Computer interface systems. *Computational Intelligence and Neuroscience*, 2019, 1–12. <https://doi.org/10.1155/2019/5427154>
- Gutierrez-Martin, L., Ongil, C. L., Lanza-Gutierrez, J. M., & Calero, J. a. M. (2024). Improving Emotion Recognition Accuracy with Personalized Clustering. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2410.03696>
- Kabir, M. H., Akhtar, N. I., Tasnim, N., Miah, A. S. M., Lee, H., Jang, S., & Shin, J. (2024). Exploring feature selection and classification techniques to improve the performance of an Electroencephalography-Based Motor Imagery Brain–Computer Interface system. *Sensors*, 24(15), 4989. <https://doi.org/10.3390/s24154989>

Multivariate Determinants of Playability in Massively Multiplayer Online Game Teams: The Case of China

Xiaoxue Gong, Lili Nurliyana Binti Abdullah, Azrul Hazri Bin Jantan, Noris Binti Mohd Norowi
Faculty of Computer Science and Information Technology, Universiti Putra Malaysia
 gs64174@student.upm.edu.my, liyana@upm.edu.my, azrulhazri@upm.edu.my,
 noris@upm.edu.my

With the advancement of computer technology, video games have evolved into platforms for human-computer interaction (HCI). Massively multiplayer online games (MMOGs) now enable thousands of players to interact in real time, creating valuable contexts for studying user experience and group behaviour (Linares et al., 2021). The global MMOG market, valued at 48.98 billion USD in 2022, is expected to reach 134.68 billion USD by 2031. China, the largest market with over 400 million users, provides abundant data for HCI and behavioural research (Ye, 2023). However, most existing research isolates social, technical, or gameplay variables, leading to fragmented findings (Cheng et al., 2019). While studies have explored social ties, communication structures, and emotional engagement (Ghawi et al., 2021; Ma et al., 2023; Zeng, 2024), no integrated model yet accounts for both player and technical factors in MMOG team play.

Current MMOG design frameworks remain insufficient to address the complex mix of social, technical, and emotional dynamics in large-scale online environments. Although trust and cooperation are crucial (Tu et al., 2023), player insights rarely inform real-time design strategies. Existing models often focus on isolated elements such as communication or leadership and neglect adaptability to players' emotional or cultural contexts (Yu & Cardoso-Leite, 2023). Likewise, established heuristic tools like HEP (Desurvire et al., 2004), GameFlow (Sweetser, 2020), and Nielsen's heuristics (Nielsen, 1994) cater mainly to single-player or simple multiplayer experiences, overlooking emergent teamwork and live interaction (Aker et al., 2020). This highlights the need for a specialised heuristic evaluation framework to assess playability, usability, and social cooperation in complex MMOG team environments. The specific details of the problem statement, research problem, and objective for this study are shown in Table 1.

Table 1. Research Question and Objective

No	Problem Statement (PS)	Research Question (RQ)	Research Objective (RO)
1	Lacks a comprehensive model that accounts for the interplay between technical design and team dynamics, as existing studies tend to isolate individual factors and overlook collaborative mechanisms essential for team playability (Jiang et al., 2023).	What are the key social, technical, and gameplay-related determinants influencing team playability in Massively Multiplayer Online Games (MMOGs)? How can these be systematically integrated into the model to support effective team collaboration and design evaluation?	To identify the multivariable determinants from social, technical, and gameplay-related factors that influence team playability in MMOGs and to conceptualize these within the team playability model.

2	Fail to integrate technical design and team dynamics such as communication, trust-building, and collaboration, highlighting the need for a comprehensive solution to evaluate and optimise team playability in multiplayer environments (Yang & Cho, 2022).	What types of design guidelines can be derived from multivariable determinants to enhance social interaction and team collaboration among MMOG players, as evaluated through the team playability model?	To propose a comprehensive set of design guidelines derived from multivariable determinants to provide game developers with actionable understandings to enhance collaborative gameplay experiences and foster more engaging team dynamics.
3	Lack of playability heuristic evaluation tools suitable for team collaboration (Aker et al., 2020).	What are the essential components of a heuristic evaluation framework tailored to team playability in MMOGs, and how effective is it in assessing cooperative gameplay?	To develop Heuristics Evaluation for Team Playability (HETP), a specialized set of evaluation criteria designed to assess teamwork, coordination, and shared engagement in cooperative MMOGs to enhance the multiplayer experience.

This study examines complex multivariable relationships that cannot be addressed by a single method. An explanatory sequential mixed-methods design will therefore be adopted, beginning with quantitative research followed by qualitative analysis for interpretation and expansion. This approach enables deeper understanding of underlying mechanisms, allows methodological flexibility, and strengthens the validity of findings through complementary qualitative insights (Creswell & Guetterman, 2024).

As illustrated in Figure 1, this study adopts an explanatory sequential mixed- methods design that begins with quantitative data collection and analysis, followed by qualitative exploration to elaborate and extend the statistical findings. This two-phase design enables the researcher to first identify patterns and relationships through empirical data and subsequently interpret the underlying causes and contextual factors in depth.

Building on the logic of this sequential design, the overall methodological process was operationalised through a structured set of research activities. Figure 2 below presents this comprehensive research process, translating the conceptual design into practical stages of implementation. Specifically, it outlines how the quantitative and qualitative components are integrated across three major stages: (1) the development of the POITP Model to identify determinants of playability, (2) the validation of design guidelines through expert and experimental evaluation, and (3) the formulation of the HETP framework to enhance team collaboration and playability evaluation in MMOGs.

In this way, Figure 1 demonstrates the theoretical foundation of the research design, while Figure 2 visualises its step-by-step execution, ensuring coherence between methodological structure and procedural implementation. This study was approved by the Institutional Review Board (JKEUPM) and adhered to ethical research standards involving human participants. Informed consent was obtained through a form outlining the study’s purpose, procedures, risks, benefits, and confidentiality measures. The consent form and participant information were presented at the start of the questionnaire, and continuation indicated voluntary participation. All responses were

anonymous, with no sensitive or identifying questions included. Participants were informed about data handling and confidentiality, and all data were securely stored and used solely for this research to protect participants' privacy and rights (Facca et al., 2020).

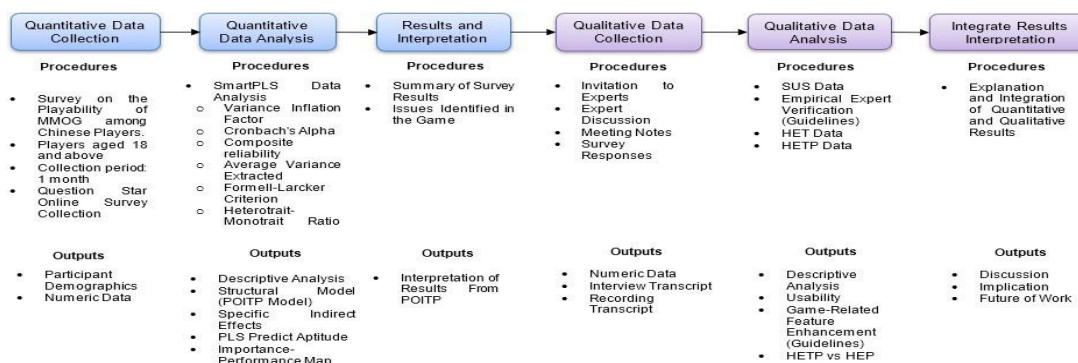


Figure 1. Explanatory Sequential Design in Mixed Method Research Flow

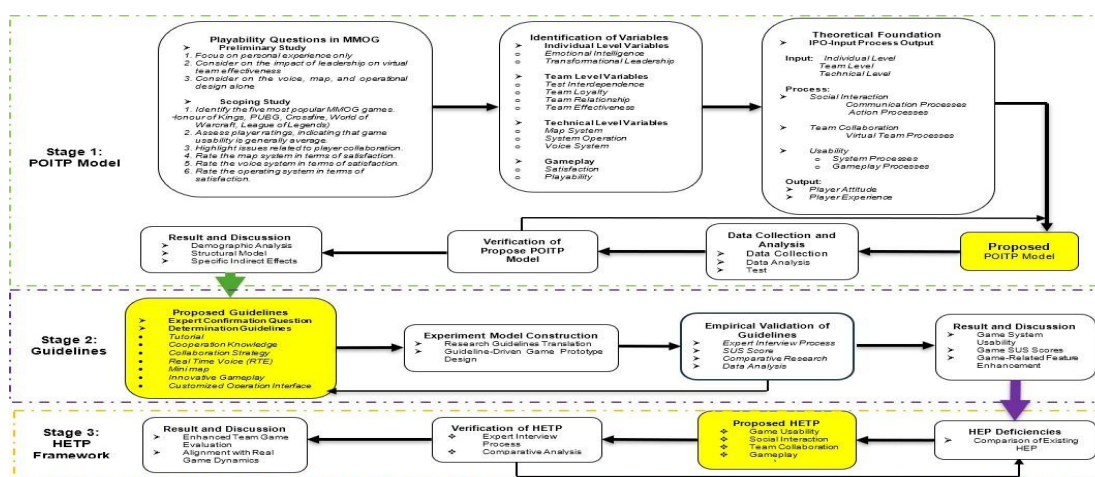


Figure 2. Research Process

References

- Aker, Ç., Rızvanoğlu, K., & Bostan, B. (2020). *Methodological review of playability heuristics*.
- Creswell, J. W., & Guetterman, T. C. (2024). Educational research: Planning, conducting, and evaluating quantitative and qualitative research, 7th edition. In *Pearson*. Pearson.
- Desurvire, H., Caplan, M., & Toth, J. A. (2004). Using heuristics to evaluate the playability of games. *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, 1509– 1512. <https://doi.org/10.1145/985921.986102>
- Facca D., Smith M. J., Shelley J., Lizotte D., & Donelle L. (2020). Exploring the ethical issues in research using digital data collection strategies with minors: A scoping review. *PLOS ONE*, 15(8), e0237875. <https://doi.org/10.1371/journal.pone.0237875>

- Ghawi, R., Müller, S., & Pfeffer, J. (2021). Improving team performance prediction in MMOGs with temporal communication networks. *Social Network Analysis and Mining*, 11(1), 65. <https://doi.org/10.1007/s13278-021-00775-7>
- Jiang, J., Wong, P. P. W., & Wang, L. (2023). Influence of Playability on Perceived Game Software Quality: A Comparison Between Chinese and South Korean Gamers. *SAGE Open*, 13(3), 21582440231197506. <https://doi.org/10.1177/21582440231197506>
- Linares, M., Gallego, M. D., & Bueno, S. (2021). Proposing a TAM-SDT-Based Model to Examine the User Acceptance of Massively Multiplayer Online Games. *International Journal of Environmental Research and Public Health*, 18(7), 3687. <https://doi.org/10.3390/ijerph18073687>
- Ma, R., Li, Y., & Kou, Y. (2023). Transparency, Fairness, and Coping: How Players Experience Moderation in Multiplayer Online Games. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–21. <https://doi.org/10.1145/3544548.3581097>
- Nielsen, J. (1994). *Enhancing the explanatory power of usability heuristics*. pp.152-158. <https://dl.acm.org/doi/abs/10.1145/191666.191729>
- Sweetser, P. (2020). GameFlow 2020: 15 Years of a Model of Player Enjoyment. In Ahmadpour N., Leong T., Ploderer B., Parker C., Webber S., Munoz D., Loke L., & Tomitsch M. (Eds), *ACM Int. Conf. Proc. Ser.* (pp. 705–711). Association for Computing Machinery; Scopus. <https://doi.org/10.1145/3441000.3441048>
- Tu, C., Tunggal, J., & Brown, S. (2023). Character immersion in video games as a form of acting. *Psychology of Popular Media*, 12(4), 405–413. <https://doi.org/10.1037/ppm0000435>
- Yang, D., & Cho, D.-M. (2022). Research on the Influence of Curiosity on MMORPG Grinding Player Experience. *Journal of Multimedia Information System*, 9(2), 127–136. <https://doi.org/10.33851/JMIS.2022.9.2.127>
- Ye, J. (2023, July 27). China's video game player population rises to a record at end-June. *Reuters*. <https://www.reuters.com/world/china/chinas-video-game-player-population-rises-record-end-june-2023-07-27/>
- Yu, H., & Cardoso-Leite, P. (2023). Video Games to Study and Improve Collaboration Skills. *Companion Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, 149–154. <https://doi.org/10.1145/3573382.3616091>
- Zeng, Y. (2024). *Learning social sequential decision making in online games* (PQDT:90793618)

Designing Co-Located Mixed Reality Gamelan Collaboration Using Non-verbal Cues

Mohamad Faris Hakimin, Noris Mohd Norowi
Universiti Putra Malaysia

Introduction

Background of the study

The advancement of immersive technologies has transformed how music is created, performed, and experienced. Virtual reality (VR) platforms enable new forms of interactive spatial music-making, allowing users to engage with instruments and sound environments in three-dimensional spaces. Applications such as *TheWaveVR* and *SynthVR* demonstrate how users can compose and perform music in virtual settings (Hamilton, 2021; Zhang et al., 2020). However, these systems primarily reflect Western music traditions, leaving traditional ensembles like Gamelan Melayu underrepresented in virtual and mixed reality (MR) environments (Tan, 2020; Dobrian, 2019).

Gamelan Melayu is a traditional ensemble music of the Malay world, featuring bronze and wooden percussion instruments, gongs, and drums. Its performance relies heavily on non-verbal communication such as eye contact, gestures, and body movements to maintain ensemble coordination and expressiveness. However, in co-located MR settings where one performer uses a head-mounted display (HMD), these cues are often obstructed by the headset, reducing visibility of facial and eye signals. This loss of communicative expressivity can disrupt synchrony, weaken social presence, and hinder collaborative flow.

This research investigates how non-verbal communication can be restored and enhanced in co-located MR environments for Gamelan musicians. The aim is to develop a design framework that supports social presence, coordination, and expressive collaboration between HMD and non-HMD performers.

Problem Statement

Despite progress in MR technology, three persistent challenges limit its effectiveness for traditional ensemble collaboration:

- 1. Loss of non-verbal cues:**

HMDs obscure facial expressions, gestures, and gaze, reducing social presence and natural coordination between users (Chen et al., 2022; Lin et al., 2022).

- 2. Lack of shared presence:**

Differences between HMD and non-HMD perspectives hinder mutual awareness, affecting timing, rhythm, and collaboration flow (Kiyokawa et al., 2022; Paulsen et al., 2024).

Addressing these challenges is essential to support meaningful, culturally grounded, and expressive collaboration in co-located MR Gamelan performance.

Aim and Research Objectives

To develop and evaluate a co-located Mixed Reality (MR) Gamelan system that supports expressive and socially present collaboration between HMD and non-HMD players through non-verbal cues, two-way interaction, and responsive audio feedback.

1. To design, a button-based interaction mapping system that restores non-verbal cues (eg., attention signals, gestures) between performers.
2. To develop a prototype MR environment that supports two-way interaction and expressive feedback.
3. To evaluate the system's effectiveness in improving social presence and performance coordination.

Research Questions

1. How can button-based interaction mapping be used to restore non-verbal cues in co-located MR environments involving HMD and non-HMD users?
2. What design strategies can enable effective two-way interaction and collaboration in MR-based Gamelan performance?
3. How does restoring non-verbal communication influence social presence and ensemble coordination?

Scopes of the Study

This research focuses on interaction between an HMD user and nearby non-HMD players in a co-located MR Gamelan setting.

The system will utilize a Meta Quest 3 MR headset with controllers to implement button-based cue mapping and visual feedback. Authentic Gamelan sound samples will create a realistic context; however, audio fidelity and spatial sound processing are not the focus. The main emphasis is on interaction design, social presence, and communication efficiency between participants.

Methodology

This study adopts a User-Centered Design (UCD) approach, emphasizing iterative prototyping, user testing, and feedback integration to ensure that the MR system aligns with performer needs and cultural context (Graham et al., 2019).

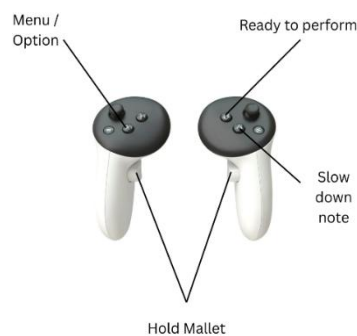
Phase 0: Preliminary Study

- Aim: Identify user needs and challenges in MR Gamelan collaboration.

- Method: Semi-structured interviews with 4–6 professional or semi-professional Gamelan players (e.g., Gangspura, Rhythm in Bronze).
- Analysis: Thematic analysis (Braun & Clarke, 2006) to extract key communication cues and usability expectations.

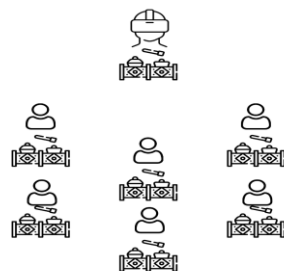
Phase 1: Restoring Non-Verbal Cues

- Aim: Design and test a button-mapped cue system to restore non-verbal communication between HMD and non-HMD players.
- Evaluation: Measure cue clarity, response timing, and user satisfaction through observation and usability questionnaires (Brooke, 1996).



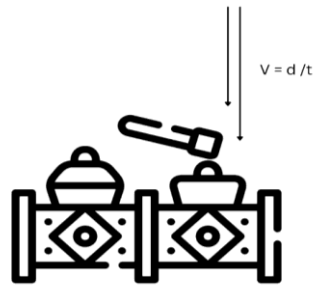
Phase 2: Co-Located Interaction and Social Presence

- Aim: Develop a two-way co-located MR interaction system using passthrough visuals and shared display feedback.
- Evaluation: Presence Questionnaire (Witmer & Singer, 1998) and qualitative interviews to assess collaboration flow and mutual awareness.



Phase 3: Enhancing Expressive Audio Responsiveness

- Aim: Implement and test velocity-sensitive audio interaction for more expressive MR Gamelan play.
- Evaluation: Observation and interview data to assess perceived realism,



expressivity, and engagement.

References

- Ahmad Saffian, K., Mohd Norowi, N., Abdullah, L. N., Sulaiman, P. S., & Musib, A. F. (2022). *Playing Gamelan Bonang in the Air: User requirements for designing a digital musical instrument for the Malay Bonang*. *Malaysian Journal of Music*, 11(1), 68–83.
- Hamilton, R. (2021). *Composing in virtual space: Immersive music interaction in XR*. Proceedings of the International Conference on New Interfaces for Musical Expression (NIME).
- Mori, K., De La Rubia, E., Guedes, A., Dooley, J., & Timmers, R. (2023). *Wish you were here: Mental and physiological effects of remote music collaboration in mixed reality*. arXiv.
- Permana, F., Tolle, H., Utaminingrum, F., & Dermawi, R. (2019). *Development of augmented reality gamelan simulation with leap motion control*. *International Journal of Interactive Mobile Technologies*, 13(12), 120–135.
- Witmer, B. G., & Singer, M. J. (1998). *Measuring presence in virtual environments: A presence questionnaire*. *Presence: Teleoperators and Virtual Environments*, 7(3), 225



About Fusion

The 7th National Symposium on Human-Computer Interaction (FUSION 2025), held on 25-26 October 2025 is organized by the Association of Human Computer Interaction Malaysia (myHCI- UX).

FUSION is a yearly symposium held in Malaysia, bringing together academicians, researchers, and practitioners in the field of Human Computer Interaction (HCI). In the context of our country that is diverse and has rich cultural (hence; fusion) the aim of this symposium is to learn from each other's ongoing research on HCI of various domains, and explore the possibilities of collaborative work.

Ultimately, the outcome of the symposium is to set a path on how HCI can contribute and make an impact in the landscape of Malaysia's ICT and digital transformation.



<http://myhci-ux.org>



myHCI-UX:Kuala Lumpur ACM SIGCHI Chapter

