Adela Coman
Simona Vasilache (Eds.)

LNCS 15786

Social Computing and Social Media

17th International Conference, SCSM 2025
Held as Part of the 27th HCI International Conference, HCII 2025
Gothenburg, Sweden, June 22–27, 2025
Proceedings, Part I

Part I

Springer

Lecture Notes in Computer Science **15786**

Founding Editors

Gerhard Goos
Juris Hartmanis

Editorial Board Members

Elisa Bertino, *Purdue University, West Lafayette, IN, USA*
Wen Gao, *Peking University, Beijing, China*
Bernhard Steffen, *TU Dortmund University, Dortmund, Germany*
Moti Yung, *Columbia University, New York, NY, USA*

The series Lecture Notes in Computer Science (LNCS), including its subseries Lecture Notes in Artificial Intelligence (LNAI) and Lecture Notes in Bioinformatics (LNBI), has established itself as a medium for the publication of new developments in computer science and information technology research, teaching, and education.

LNCS enjoys close cooperation with the computer science R & D community, the series counts many renowned academics among its volume editors and paper authors, and collaborates with prestigious societies. Its mission is to serve this international community by providing an invaluable service, mainly focused on the publication of conference and workshop proceedings and postproceedings. LNCS commenced publication in 1973.

Adela Coman · Simona Vasilache
Editors

Social Computing and Social Media

17th International Conference, SCSM 2025
Held as Part of the 27th HCI International Conference, HCII 2025
Gothenburg, Sweden, June 22–27, 2025
Proceedings, Part I

Springer

Editors
Adela Coman
University of Bucharest
Bucharest, Romania

Simona Vasilache
University of Tsukuba
Tsukuba, Japan

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Computer Science
ISBN 978-3-031-93538-1 ISBN 978-3-031-93539-8 (eBook)
https://doi.org/10.1007/978-3-031-93539-8

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Switzerland AG 2025

This work is subject to copyright. All rights are solely and exclusively licensed 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, expressed 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.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

If disposing of this product, please recycle the paper.

Foreword

The HCI International (HCII) conference was founded in 1984 by Gavriel Salvendy (Purdue University, USA, Tsinghua University, P.R. China, and University of Central Florida, USA) and the first event of the series, "1st USA-Japan Conference on Human-Computer Interaction", was held in Honolulu, Hawaii, USA, 18–20 August. Since then, HCI International is held jointly with several Thematic Areas and Affiliated Conferences, with each one under the auspices of a distinguished international Program Board and under one management and one registration. Twenty-seven HCI International Conferences have been organized so far (every two years until 2013, and annually thereafter).

Last year, we celebrated 40 years since the establishment of the HCII conference, which has been a hub for presenting groundbreaking research and novel ideas and collaboration for people from all over the world. Over the years, this conference has served as a platform for scholars, researchers, industry experts, and students to exchange ideas, connect, and address challenges in the ever-evolving HCI field. The conference has evolved itself, adapting to new technologies and emerging trends, while staying committed to its core mission of advancing knowledge and driving change.

The 27th International Conference on Human-Computer Interaction, HCI International 2025 (HCII 2025), was held as an 'on-site' conference at the Gothia Towers Hotel and Swedish Exhibition & Congress Centre, in Gothenburg, Sweden, on June 22–27, 2025, with the additional option for 'on-line' participation. It incorporated the 21 thematic areas and affiliated conferences listed below.

A total of 7972 individuals from academia, research institutes, industry, and government agencies from 92 countries submitted contributions. 1430 papers and 355 posters (as short research papers) are included in the volumes of the proceedings published just before the start of the conference, and which are listed below. The contributions thoroughly cover the entire field of human-computer interaction, highlight the evolving role of computers in diverse contexts, and demonstrate how HCI research is shaping and improving user experiences across a wide range of domains, influencing technological progress and its effective integration into various sectors.

The HCII conference also offers the option of presenting 'Late Breaking Work', both for papers and posters, with the corresponding proceedings volumes published after the conference. Full papers are included in the 'HCII 2025 - Late Breaking Papers' volumes of the proceedings published in the Springer LNCS series, while 'Poster Extended Abstracts' are included as short research papers in the 'HCII 2025 - Late Breaking Posters' volumes published in the Springer CCIS series.

I would like to thank the Program Board Chairs and the members of the Program Boards of all thematic areas and affiliated conferences for their contribution towards the high scientific quality and overall success of the HCI International 2025 conference. Their manifold support including paper reviews (via a single-blind review process, with a minimum of two reviews per submission), session organization, and their willingness to act as goodwill ambassadors for the conference is most highly appreciated.

This conference would not have been possible without the continuous and unwavering support and advice of Gavriel Salvendy, founder, General Chair Emeritus, and Scientific Advisor. For his outstanding efforts, I would like to express my sincere appreciation to Abbas Moallem, Communications Chair and Editor of HCI International News.

June 2025 Constantine Stephanidis

HCI International 2025 Thematic Areas and Affiliated Conferences

- HCI: Human-Computer Interaction Thematic Area
- HIMI: Human Interface and the Management of Information Thematic Area
- EPCE: 22nd International Conference on Engineering Psychology and Cognitive Ergonomics
- AC: 19th International Conference on Augmented Cognition
- UAHCI: 19th International Conference on Universal Access in Human-Computer Interaction
- CCD: 17th International Conference on Cross-Cultural Design
- SCSM: 17th International Conference on Social Computing and Social Media
- VAMR: 17th International Conference on Virtual, Augmented and Mixed Reality
- DHM: 16th International Conference on Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management
- DUXU: 14th International Conference on Design, User Experience, and Usability
- C&C: 13th International Conference on Culture and Computing
- DAPI: 13th International Conference on Distributed, Ambient and Pervasive Interactions
- HCIBGO: 12th International Conference on HCI in Business, Government and Organizations
- LCT: 12th International Conference on Learning and Collaboration Technologies
- ITAP: 11th International Conference on Human Aspects of IT for the Aged Population
- AIS: 7th International Conference on Adaptive Instructional Systems
- HCI-CPT: 7th International Conference on HCI for Cybersecurity, Privacy and Trust
- HCI-Games: 7th International Conference on HCI in Games
- MobiTAS: 7th International Conference on HCI in Mobility, Transport and Automotive Systems
- AI-HCI: 6th International Conference on Artificial Intelligence in HCI
- MOBILE: 6th International Conference on Human-Centered Design, Operation and Evaluation of Mobile Communications

List of Conference Proceedings Volumes Appearing Before the Conference

1. LNCS 15766, Human-Computer Interaction - Part I, edited by Masaaki Kurosu and Ayako Hashizume
2. LNCS 15767, Human-Computer Interaction - Part II, edited by Masaaki Kurosu and Ayako Hashizume
3. LNCS 15768, Human-Computer Interaction - Part III, edited by Masaaki Kurosu and Ayako Hashizume
4. LNCS 15769, Human-Computer Interaction - Part IV, edited by Masaaki Kurosu and Ayako Hashizume
5. LNCS 15770, Human-Computer Interaction - Part V, edited by Masaaki Kurosu and Ayako Hashizume
6. LNCS 15771, Human-Computer Interaction - Part VI, edited by Masaaki Kurosu and Ayako Hashizume
7. LNCS 15772, Human-Computer Interaction - Part VII, edited by Masaaki Kurosu and Ayako Hashizume
8. LNCS 15773, Human Interface and the Management of Information: Part I, edited by Hirohiko Mori and Yumi Asahi
9. LNCS 15774, Human Interface and the Management of Information: Part II, edited by Hirohiko Mori and Yumi Asahi
10. LNCS 15775, Human Interface and the Management of Information: Part III, edited by Hirohiko Mori and Yumi Asahi
11. LNAI 15776, Engineering Psychology and Cognitive Ergonomics: Part I, edited by Don Harris and Wen-Chin Li
12. LNAI 15777, Engineering Psychology and Cognitive Ergonomics: Part II, edited by Don Harris and Wen-Chin Li
13. LNAI 15778, Augmented Cognition: Part I, edited by Dylan D. Schmorrow and Cali M. Fidopiastis
14. LNAI 15779, Augmented Cognition: Part II, edited by Dylan D. Schmorrow and Cali M. Fidopiastis
15. LNCS 15780, Universal Access in Human-Computer Interaction: Part I, edited by Margherita Antona and Constantine Stephanidis
16. LNCS 15781, Universal Access in Human-Computer Interaction: Part II, edited by Margherita Antona and Constantine Stephanidis
17. LNCS 15782, Cross-Cultural Design: Part I, edited by Pei-Luen Patrick Rau
18. LNCS 15783, Cross-Cultural Design: Part II, edited by Pei-Luen Patrick Rau
19. LNCS 15784, Cross-Cultural Design: Part III, edited by Pei-Luen Patrick Rau
20. LNCS 15785, Cross-Cultural Design: Part IV, edited by Pei-Luen Patrick Rau

21. LNCS 15786, Social Computing and Social Media: Part I, edited by Adela Coman and Simona Vasilache
22. LNCS 15787, Social Computing and Social Media: Part II, edited by Adela Coman and Simona Vasilache
23. LNCS 15788, Virtual, Augmented and Mixed Reality: Part I, edited by Jessie Y. C. Chen and Gino Fragomeni
24. LNCS 15789, Virtual, Augmented and Mixed Reality: Part II, edited by Jessie Y. C. Chen and Gino Fragomeni
25. LNCS 15790, Virtual, Augmented and Mixed Reality: Part III, edited by Jessie Y. C. Chen and Gino Fragomeni
26. LNCS 15791, Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management: Part I, edited by Vincent G. Duffy
27. LNCS 15792, Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management: Part II, edited by Vincent G. Duffy
28. LNCS 15793, Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management: Part III, edited by Vincent G. Duffy
29. LNCS 15794, Design, User Experience, and Usability: Part I, edited by Martin Schrepp
30. LNCS 15795, Design, User Experience, and Usability: Part II, edited by Martin Schrepp
31. LNCS 15796, Design, User Experience, and Usability: Part III, edited by Martin Schrepp
32. LNCS 15797, Design, User Experience, and Usability: Part IV, edited by Martin Schrepp
33. LNCS 15798, Design, User Experience, and Usability: Part V, edited by Martin Schrepp
34. LNCS 15799, Design, User Experience, and Usability: Part VI, edited by Martin Schrepp
35. LNCS 15800, Culture and Computing: Part I, edited by Matthias Rauterberg
36. LNCS 15801, Culture and Computing: Part II, edited by Matthias Rauterberg
37. LNCS 15802, Distributed, Ambient and Pervasive Interactions: Part I, edited by Norbert A. Streitz and Shin'ichi Konomi
38. LNCS 15803, Distributed, Ambient and Pervasive Interactions: Part II, edited by Norbert A. Streitz and Shin'ichi Konomi
39. LNCS 15804, HCI in Business, Government and Organizations: Part I, edited by Fiona Fui-Hoon Nah and Keng Leng Siau
40. LNCS 15805, HCI in Business, Government and Organizations: Part II, edited by Fiona Fui-Hoon Nah and Keng Leng Siau
41. LNCS 15806, Learning and Collaboration Technologies: Part I, edited by Brian K. Smith and Marcela Borge
42. LNCS 15807, Learning and Collaboration Technologies: Part II, edited by Brian K. Smith and Marcela Borge
43. LNCS 15808, Learning and Collaboration Technologies: Part III, edited by Brian K. Smith and Marcela Borge

44. LNCS 15809, Human Aspects of IT for the Aged Population: Part I, edited by Qin Gao and Jia Zhou
45. LNCS 15810, Human Aspects of IT for the Aged Population: Part II, edited by Qin Gao and Jia Zhou
46. LNCS 15811, Human Aspects of IT for the Aged Population: Part III, edited by Qin Gao and Jia Zhou
47. LNCS 15812, Adaptive Instructional System: Part I, edited by Robert A. Sottilare and Jessica Schwarz
48. LNCS 15813, Adaptive Instructional System: Part II, edited by Robert A. Sottilare and Jessica Schwarz
49. LNCS 15814, HCI for Cybersecurity, Privacy and Trust: Part I, edited by Abbas Moallem
50. LNCS 15815, HCI for Cybersecurity, Privacy and Trust: Part II, edited by Abbas Moallem
51. LNCS 15816, HCI in Games, edited by Xiaowen Fang
52. LNCS 15817, HCI in Mobility, Transport and Automotive Systems: Part I, edited by Heidi Krömker
53. LNCS 15818, HCI in Mobility, Transport and Automotive Systems: Part II, edited by Heidi Krömker
54. LNAI 15819, Artificial Intelligence in HCI: Part I, edited by Helmut Degen and Stavroula Ntoa
55. LNAI 15820, Artificial Intelligence in HCI: Part II, edited by Helmut Degen and Stavroula Ntoa
56. LNAI 15821, Artificial Intelligence in HCI: Part III, edited by Helmut Degen and Stavroula Ntoa
57. LNAI 15822, Artificial Intelligence in HCI: Part IV, edited by Helmut Degen and Stavroula Ntoa
58. LNCS 15823, Human-Centered Design, Operation and Evaluation of Mobile Communications: Part I, edited by June Wei and George Margetis
59. LNCS 15824, Human-Centered Design, Operation and Evaluation of Mobile Communications: Part II, edited by June Wei and George Margetis
60. CCIS 2522, HCI International 2025 Posters - Part I, edited by Constantine Stephanidis, Margherita Antona, Stavroula Ntoa and Gavriel Salvendy
61. CCIS 2523, HCI International 2025 Posters - Part II, edited by Constantine Stephanidis, Margherita Antona, Stavroula Ntoa and Gavriel Salvendy
62. CCIS 2524, HCI International 2025 Posters - Part III, edited by Constantine Stephanidis, Margherita Antona, Stavroula Ntoa and Gavriel Salvendy
63. CCIS 2525, HCI International 2025 Posters - Part IV, edited by Constantine Stephanidis, Margherita Antona, Stavroula Ntoa and Gavriel Salvendy
64. CCIS 2526, HCI International 2025 Posters - Part V, edited by Constantine Stephanidis, Margherita Antona, Stavroula Ntoa and Gavriel Salvendy
65. CCIS 2527, HCI International 2025 Posters - Part VI, edited by Constantine Stephanidis, Margherita Antona, Stavroula Ntoa and Gavriel Salvendy

66. CCIS 2528, HCI International 2025 Posters - Part VII, edited by Constantine Stephanidis, Margherita Antona, Stavroula Ntoa and Gavriel Salvendy
67. CCIS 2529, HCI International 2025 Posters - Part VIII, edited by Constantine Stephanidis, Margherita Antona, Stavroula Ntoa and Gavriel Salvendy

https://2025.hci.international/proceedings

Preface

The 17th International Conference on Social Computing and Social Media (SCSM 2025) was an affiliated conference of the HCI International (HCII) conference. It provided an established international forum for the exchange and dissemination of scientific information related to social computing and social media, addressing a broad spectrum of issues expanding our understanding of current and future issues in these areas. The conference welcomed qualitative and quantitative research papers on a diverse range of topics related to the design, development, assessment, use, and impact of social media.

A considerable number of papers this year focused on research on the design, development, and evaluation of social media, also delving into aspects related to user experience and user behavior. The undeniable influence of Artificial Intelligence on the technological landscape prompted numerous works focused on the use of AI and Language Models in social media, investigating their multifaceted impact on the field. The power of social media and its positive impact across various application domains inspired contributions regarding education and learning, culture, business, and eCommerce. In the context of learning, the topics explored included academic writing, learning experience, and the impact of AI in education-related social media and platforms. As editors of these SCSM proceedings volumes, we are pleased to present this unique and diverse compilation of topics offering valuable insights and advancing our understanding of the current and future issues in the field.

Two volumes of the HCII 2025 proceedings are dedicated to this year's edition of the SCSM conference. The first volume focuses on topics related to Designing and Developing Social Interactions, Social Media in Learning and Education, and Supporting Communication and Psychological Well-Being. The second volume focuses on topics related to User Behavior and Experience in Social Media, and AI and Social Network Analysis.

The papers in these volumes were accepted for publication after a minimum of two single-blind reviews from the members of the SCSM Program Board or, in some cases, from members of the Program Boards of other affiliated conferences. We would like to thank all of them for their invaluable contribution, support, and efforts.

June 2025

Adela Coman
Simona Vasilache

17th International Conference on Social Computing and Social Media (SCSM 2025)

Program Board Chairs: **Adela Coman,** *University of Bucharest, Romania*, and **Simona Vasilache**, *University of Tsukuba, Japan*

Francisco Javier Alvarez-Rodriguez, *Universidad Autónoma de Aguascalientes, Mexico*
Andria Andriuzzi, *Université Jean Monnet, France*
Karine Berthelot-Guiet, *Sorbonne University, France*
James Braman, *Community College of Baltimore County, USA*
Magdalena Brzezinska, *WSB Merito University, Poland*
Adheesh Budree, *University of Cape Town, South Africa*
Hung-Hsuan Huang, *University of Fukuchiyama, Japan*
Ajrina Hysaj, *University of Wollongong in Dubai, United Arab Emirates*
Ayaka Ito, *Reitaku University, Japan*
Jeannie Lee, *Singapore Institute of Technology, Singapore*
Gabriele Meiselwitz, *Towson University, USA*
Takashi Namatame, *Chuo University, Japan*
Harry Nguyen, *University College Cork, Ireland*
Kohei Otake, *Sophia University, Japan*
Daniela Quinones, *Pontificia Universidad Católica de Valparaíso, Chile*
Virginica Rusu, *Universidad de Playa Ancha, Chile*
Cristian Rusu, *Pontificia Universidad Católica de Valparaíso, Chile*
Tomislav Stipancic, *University of Zagreb, Croatia*
Yuanqiong Wang, *Towson University, USA*

The full list with the Program Board Chairs and the members of the Program Boards of all thematic areas and affiliated conferences of HCII 2025 is available online at:

http://www.hci.international/board-members-2025.php

HCI International 2026 Conference

The 28th International Conference on Human-Computer Interaction, HCI International 2026, will be held jointly with the affiliated conferences at the Montréal Convention Centre (Palais des congrès de Montréal), in Montreal, Canada, 26–31 July 2026. It will cover a broad spectrum of themes related to Human-Computer Interaction, including theoretical issues, methods, tools, processes, and case studies in HCI design, as well as novel interaction techniques, interfaces, and applications. The proceedings will be published by Springer (part of Springer Nature) in a multi-volume set. More information will become available on the conference website: https://2026.hci.international/.

General Chair
Prof. Constantine Stephanidis
University of Crete and ICS-FORTH
Heraklion, Crete, Greece
Email: general_chair@2026.hci.international

https://2026.hci.international/

Contents – Part I

Designing and Developing Social Interactions

Key Factors for the Effective Development and Implementation of Geospatial Dashboards for Data-Driven Decision Making 3
Oscar Ancán-Bastías, Carlos Cares, Jesús Peral Cortés, and Jaime Díaz-Arancibia

A Serious Game Approach for Teaching Requirements Engineering: User Experience Evaluation 20
Tomás Castillo, Benjamín Chamorro, Ignacio Catalán, Pablo Schwarzenberg, Luis Rojas, and Fei Luo

Identification of Older Adults' Characteristics that Affect the Usability of Mobile Applications: A Tertiary Study 37
John W. Castro, Agustín I. Astudillo, and Luis Rojas

Generating Product Descriptions Using Customer Reviews on E-Commerce Sites 57
Aina Ishikawa, Joshujio Takanami, and Kohei Otake

Improving Intention Recognition Efficiency: A Study on Skeletal Data Dimensionality Reduction and Neural Architectures 70
Luka Orsag, Tomislav Stipancic, Leon Koren, and Matija Zidaric

First Steps Toward the Agile Integration of Information Architecture into a User-Centered Development Process 87
Luis Rojas, Katrina Sorbello, Vicente Marin, Fei Luo, John W. Castro, Avon Collis, and Pablo Schwarzenberg

LLM-Driven Augmented Reality Puppeteer: Controller-Free Voice-Commanded Robot Teleoperation 97
Yuchong Zhang, Bastian Orthmann, Michael C. Welle, Jonne Van Haastregt, and Danica Kragic

Social Media in Learning and Education

Redimensioning Visible Learning and Teaching in the Dynamics of a New Reality 115
Rudina Alimerko and Deniz Çupi Xhoga

Generative AI in Education: Exploring EAP Faculty Perspectives at a Multicultural UAE University 125
Asim Bashir, Saima Nomaan, and Ajrina Hysaj

The Challenges Faced by Albanian Teachers in the Use of Media Technology During Teaching 141
Deniz Çupi and Rudina Alimerko

Undergraduate Students' Journey with AI in the United Arab Emirates 155
Georgina Farouqa, Ajrina Hysaj, Zaydoon Hatamleh, Amjad Owais, Othman Ibrahim, Laith Hiasat, and Sara Azeem Khan

Perspectives of Faculty on the Easiness and Usefulness of AI Tutoring Systems in Higher Education 172
Doaa Hamam, Yasser Ouda, Ajrina Hysaj, Mahinour Ezzat, and Sara Azeem Khan

Exploring the Use of Paraphrasing Tools in Academic Writing and Its Potential Relation with Instances of Plagiarism 180
Ajrina Hysaj, Sara Azeem Khan, and Georgina Farouqa

"Do As Your Teacher Tells You!" How Is AI Use Regulated in Nordic Higher Education Institutions? 194
Sigurbjörg Jóhannesdóttir, Päivi Kousa, Salme Korkala, Jenny Tarvainen, Kristín Björnsdóttir, Judit Háhn, Susanne Dau, Eva Irene Brooks, Jeanette Sjöberg, Skúlína Hlíf Kjartansdóttir, Nataliya Berbyuk Lindström, Sylvana Sofkova Hashemi, Mona Lundin, Shola Oyedeji, and Robin Støckert

A Property Checklist for Evaluating the Student Experience with Consideration of Cultural Aspects 219
Nicolás Matus, Cristian Rusu, and Federico Botella

Human-Robot Interaction in Higher Education: A Literature Review 236
Sebastian Matus and Sandra Cano

Artificial Intelligence in Higher Education: Student Perceptions of the Adoption and Integration in Ghana, West Africa 257
Eric Owusu, Ralitsa Debrah, and Helina Oladapo

A Management Model for Evaluating Scientific Productivity in Chilean Universities: A Case Analysis 272
María Angélica S. Piñones, María Paz O. Godoy, Geraldine Aroca, and Isidora Azocar

Virtual Reality Meets Social Media: Transforming Skill Acquisition in Physiotherapy, Veterinary Surgery, and Driver Training 290
Sutasinee Thovuttikul, Chinawat Isradisaikul, Siriporn Peansukmanee, Siriphan Kongsawasdi, and Kittichai Wantanajittikul

Addressing Students' Absenteeism in Northeastern U.S. Schools: The Role of Technological Interventions for High School Students 303
Miselta W. Tita, Johannes A. Badejo, and Joyram Chakraborty

Supporting Communication and Psychological Well-Being

Evolution of Emotional Response of PLEA—An Embodied Virtual Being with Emotional Capabilities 319
Gordana Jugo, Leon Koren, Ivan Peko, and Tomislav Stipancic

CLBoundaryManager: A System Facilitating Boundary Management and Border Crossing Between Life Domains 335
Moritz Maleck and Tom Gross

Chinese Talk About Depression on Tiktok and Its Deliberative Differences from the Russian Online Dialogue 351
Alena Rodicheva and Svetlana S. Bodrunova

Are Strength-Related Episodes Provided by Others Useful for Improving Positive Emotions? Toward the Design of a Computer-Based Self-Esteem Enhancement System 360
Tomoro Watababe, Masayuki Ando, Kouyou Otsu, and Tomoko Izumi

Telling the Truth: AI-Driven Optimization of Online Communication 375
Xinyue Xu and Ruonan Zheng

GlobalUP: Leveraging Social Media to Provide Social Support for Chinese Cultural Groups in Unfamiliar Environments 395
Haowei Xu and Xin Shu

Author Index 409

Contents – Part II

User Behavior and Experience in Social Media

Analyzing TikTok Posts from Immersive Video Experience Exhibitions Audiences: Exploring Visitors' Behavior and Experience 3
Yiwei Chen, Geoffrey Alan Rhodes, and Jianhua Yang

Exploring the Effectiveness of VTuber in Promotional Videos Through Eye-Tracking Experiments 21
Miao-Hsien Chuang and I. Gede Putu Dinanda Putra Pratama

Engagement Behaviour in Online Communities: Which Motives and Incentives Drive Customers to Participate? 31
Esther Federspiel, Matthias Baldauf, Andreas Peter, and Dorothea Schaffner

Exploring User Reactions to AI-Curated Exhibits-Emotional Engagement and Social Interaction in Digital Cultural Spaces 48
Qihan Guo, Qingshen Meng, He Li, RuiChi Li, Peng Zhang, Mingxi Shi, and Kyoungyong Lee

Alexa, Why Can't You Hear Our Accents: Cross Cultural Studies on the Inclusivity of Voice Recognition Systems 62
Yvonne Kamegne, Eric Owusu, and Helina Oladapo

Impact of Social Experience on Consumer Subscription Behavior in Paid Music Platforms: A S-O-R Model 72
Lili Liu, Jingying Wu, Panpan Wang, Yiting Zhong, Xi Feng, and Yuanna Tian

Investigating Interaction Patterns of Co-writing with AI: A Randomized Vignette Study on the Impact of Writing Perspectives 83
Teresa Luther, Michel Sommer, and Joachim Kimmerle

Evaluating the Tourist Experience on the Expedia.Com Website 95
Oriella Ortiz, Cristian Rusu, Virginica Rusu, Camila Saavedra, Ricardo Aguilera, Sergio Tapia, and Nicolás Matus

A Preliminary Property Checklist for Evaluating the Tourist Experience 109
Oriella Ortiz, Virginica Rusu, Cristian Rusu, and Nicolás Matus

Product Preference Analysis by Customer Segment Using Multiple Data Sources 123
Nonoka Takata and Takashi Namatame

Research on the Application of AIGC in Short Video Creation 135
Ouyang Wen and Hu Xin

Trends and Behavioral Risk Factors Affecting the Health Status of African Immigrants in the United States: Insights from a Cross-Sectional Study on Cardiovascular Disease 162
Nana Yeboah and Eric Owusu

Micromovement Analysis in Human-Robot Interaction for Direct Robot Manipulation 177
Matija Zidaric, Tomislav Stipancic, Leon Koren, and Luka Orsag

AI and Social Network Analysis

AI-Assisted Bias Detection of US Digital Diplomacy in Russia (2009–2023): A ChatGPT Approach 189
Mariia Bazlutckaia, Anna Sytnik, Timofey Tsvetkov, and Petr Punchenko

Detection of Depression in the Russian-Language Online Discussions: The 'AI Turn' and the Lack of Social Efficiency 210
Svetlana S. Bodrunova

Generative Artificial Intelligence in Agile Software Development Processes: A Literature Review Focused on User eXperience 228
Héctor Cornide-Reyes, Diego Monsalves, Eduardo Durán, Fabián Silva-Aravena, and Jenny Morales

Epistemic Mathematical Models for Analyzing Meta-opinions on Social Networks 247
Denis Fedyanin

Social Networks as a Driving Force for Legal Change: Emotional Interaction and the Impact of Court Broadcasts in Kazakhstan 257
Galiya Ibrayeva, Diana Tleugazina, and Aigul Ramazan

Customer Churn Prediction Model Using Gravitational Search Algorithm 272
Ryotaro Inui and Takashi Namatame

Development and Assessment of a Pricing Prediction Model for Online Flea Markets 282
Naoki Matsuo and Takashi Namatame

Automating the Investigation of Complaints on Social Networks Using Artificial Intelligence 293
Kamilla Nigmatullina and Renat Kasymov

Automated Detection of Human Values in Texts: ML Challenges and Performance Benchmarks 304
Olga Rink, Archil Maysuradze, Artem Fedorov, Roman Ischenko, Anna Korchagina, Andrey Tabachenkov, Ilya Tsybanov, and Konstantin Vorontsov

Predicting Information Foraging on Q&A Websites 322
Abim Sedhain, Sruti Srinivasa Ragavan, Brett McKinney, Shahnewaz Leon, and Sandeep Kaur Kuttal

Dynamics of COVID-19 Misinformation: An Analysis of Conspiracy Theories, Fake Remedies, and False Reports 343
Nirmalya Thakur, Mingchen Shao, Victoria Knieling, Vanessa Su, Andrew Bian, and Hongseok Jeong

Quantitative and Intensity-Based Analysis of Complaints on Parenting Q&A Sites 360
Tomoki Yoshimi and Takashi Namatame

Author Index 371

Designing and Developing Social Interactions

Key Factors for the Effective Development and Implementation of Geospatial Dashboards for Data-Driven Decision Making

Oscar Ancán-Bastías[1,2], Carlos Cares[1], Jesús Peral Cortés[2], and Jaime Díaz-Arancibia[1(✉)]

[1] Departamento de Ciencias de la Computación e Informática, Universidad de La Frontera, Temuco, Chile
jaimeignacio.diaz@ufrontera.cl

[2] Department of Software and Computing Systems, University of Alicante, 03690 Alicante, Spain

Abstract. The development of interactive analysis dashboards with georeferenced information derived from large volumes of data poses a significant challenge due to the lack of studies that establish a unified strategy for designing effective and functional dashboards. This limitation slows the development of analytical dashboards that require specific features, different from those offered by most tools available on the market. The objective of this study is to develop a procedure for implementing georeferenced analysis dashboards by identifying critical factors and the sequence in which they should be applied. Through a scoping review of the literature and subsequent content analysis, common patterns and recurring approaches in the design of georeferenced analysis dashboards were identified, providing insights aimed at systematizing the development process of these tools. As a result, the study presents a catalog of factors influencing the design and development of the foundational structure of georeferenced analysis dashboards, along with a procedure for their effective implementation. The primary contribution of this research is to provide a framework for the creation of georeferenced dashboards that support decision-making, offering development teams a catalog of key factors and an implementation sequence that facilitates the effective design of these tools, enabling the eventual automation of the process through artificial intelligence.

Keywords: Geospatial dashboards · Data-driven decision making · Spatial data visualization · Dashboard design

1 Introduction

Georeferenced analysis dashboards provide a clear and accessible visualization of spatial data, enabling decision-makers to analyze geographical patterns, iden-

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 3–19, 2025.
https://doi.org/10.1007/978-3-031-93539-8_1

tify critical areas, and optimize the distribution of resources based on specific territorial needs.

Typically, they are used in a wide range of contexts and allow for monitoring and exploring various dimensions of data [11]. One of the main characteristics of this type of technology lies in the large volumes of data from multiple sources—such as health, finance, marketing, IoT sensors, and satellite imagery—addressing challenges related to data fusion, veracity, and indicator visualization [29].

Several commercial tools facilitate the creation of interactive dashboards. Among the leading commercial tools are *ArcGIS Dashboards*, *Microsoft PowerBI*, *SAP Analytics Cloud*, and *Tableau*, among others. Additionally, there are other platforms such as *Grafana* and *Metabase*, which, while not entirely commercial, offer complementary capabilities for data visualization and analysis. However, the costs and rigidity of these platforms make it difficult to adapt the core structure of dashboards and their components, presenting obstacles for meeting specific or complex needs [27].

Nasir et al. [18] explore best practices and common principles for dashboard design, highlighting the need for a unified methodology applicable across diverse domains. Developing a universal dashboard design method remains a challenge due to the lack of a standardized approach. In this context, the importance of customization and simplicity is emphasized as fundamental principles.

Given this context, this research aims to identify common factors in the design and development of georeferenced analysis dashboards and propose an approach that facilitates their automation. Through a scoping review of the literature and subsequent content analysis, common factors and recurring approaches in the design of georeferenced analysis dashboards were identified, providing insights aimed at systematizing the development process of these tools. As a result, the study presents a catalog of factors influencing the design and development of these dashboards, as well as an algorithm for their effective implementation. These elements are suitable for automation through AI, which would further accelerate the deployment process.

This research is therefore organized as follows: Sect. 2 presents the theoretical background related to georeferenced analysis dashboards. Section 3 details the research method used. Section 4 outlines the main results of the study and their discussion, and finally, Sect. 5 presents the conclusions derived from this research.

2 Background

This section reviews the background of georeferenced analysis dashboards, focusing on their evolution, key features, and applications, followed by an examination of effective design strategies for integrating geospatial data and enhancing usability.

2.1 Georeferenced Analysis Dashboards

A georeferenced analysis dashboard is an interactive online tool that enables the visualization, analysis, and management of spatial data through the deployment of data on various types of maps. These dashboards are typically combined in a single view with charts and tabular data to facilitate the interpretation of patterns, relationships, and trends in spatial contexts, supporting location-based decision-making [19].

A fundamental aspect of these tools is georeferenced data, which consists of features, support, and attributes. The types of features include points (precise locations), lines (sequences of connected points), areas (regions bounded by lines), and volumes (three-dimensional objects with height in addition to horizontal extent). Support refers to the spatial extent and geometric properties of an entity, such as its size, shape, and orientation, while attributes represent observed or measured values associated with the entities [15]. Currently, all these elements are available in a variety of map management tools, both commercial and open-source.

Traditionally, public and private organizations have incorporated analysis dashboards as complementary tools. However, the COVID-19 pandemic highlighted the importance of understanding user experience in geospatial data visualizations beyond aspects such as speed or accuracy in specific tasks [32]. In this regard, analysis dashboards that integrate maps present unique features, as they allow users to interactively analyze spatial data, a critical aspect in applications such as smart cities, urban resource management, and public health [2].

Additionally, the use of analytical models and dynamic maps as key components for real-time visualization and pattern prediction in urban contexts has been identified to enhance smart city management through the integration of technologies [13]. For smart cities, it is crucial to integrate different types of data to generate adaptable views that allow users to adjust the analysis dashboards to their needs [8].

2.2 Design Strategies

In the context of georeferenced analysis dashboards for big data governance (BDG) in public safety services, such as disaster prevention and urban risk management, a scenario-based meta-model approach has been proposed. This approach enables the generation of customized scenarios under the Antecedent-Behavior-Consequence (ABC) theory, which has helped reduce uncertainties in user decision-making [16].

In this regard, [21] highlights that dashboards designed with a designer-driven (DD) approach excel at efficiently communicating information but offer lower interactivity. In contrast, dashboards with higher interactivity (UD) improve flow experience and user performance in critical and high-complexity environments.

In most cases, a dashboard can adequately meet functional requirements. However, non-functional requirements, such as usability, require a distinct evaluation that can determine the tool's success or failure. According to [20], although

usability evaluation methods have evolved to address the increasing complexity of interfaces, they still face specific challenges when assessing interactive components, such as maps.

The use of design patterns for usability-focused dashboards is a well-established practice aimed at improving user experience. Patterns provide reusable solutions to common design problems, facilitating consistency and user understanding [4]. In the context of map-based dashboards, these patterns must consider factors such as spatial distribution, visual hierarchy, and the ability to interact with geospatial data.

According to [14], selecting appropriate visualizations for the context is essential for communicating complex patterns. In the case of map-based dashboards, this involves correctly choosing map types, scales, and complementary visualizations that optimize data interpretation and analysis. Similarly, data organization and presentation must feel natural, particularly in maps, to avoid cognitive overload and enhance decision-making. This includes the use of clear legends, effective colors, and responsive designs [10].

The ISO 9241-210:2019 standard [12] has improved the design approach for interactive systems by establishing a set of essential activities as mandatory requirements during system development. However, implementing the standard can be challenging for organizations with limited resources or expertise. This could lead to resistance to full adoption or superficial application of the principles, reducing their impact on improving human-system interaction.

Despite advancements in georeferenced analysis dashboards, several gaps remain. User experience in geospatial visualizations is still largely assessed through basic metrics such as speed and accuracy, particularly in critical domains such as smart cities and public health. Dashboard design patterns are not sufficiently adapted to dynamic maps, lacking proper integration of visual hierarchy and spatial data interaction. The usability evaluation of interactive components, such as maps, remains challenging due to interface complexity.

3 Method

This study employs a set of methodological approaches aimed at achieving a better understanding of the object of study.

3.1 Selection and Review of Articles

This phase employed a Scoping Review based on Arksey & O'Malley [3] to broadly explore the available literature on map-based dashboard development. This approach facilitated the identification of key concepts, types of evidence, and research gaps. Indexed databases (ACM DL, IEEE Xplore, Science Direct) were consulted, and 14 relevant articles addressing the topic were collected.

3.2 Identification of Key Components

Following the initial selection of articles, a detailed review and thematic classification were conducted, adhering to the principles of qualitative analysis by Creswell [9]. During this process, the most relevant concepts were extracted and grouped according to their theoretical relevance to UX and design. As a result, a list of components or criteria was generated, recurrently referenced in academic literature and traditional usability evaluation guidelines such as [7,25,31].

3.3 Preliminary Application of the Proposal in a Pilot

Based on the stages described, a dashboard was constructed, incorporating the components in a structured manner. The design and functionality of the system were grounded in the algorithm and findings derived from the Scoping Review, ensuring coherence between scientific evidence and technical implementation. Functionality and performance tests confirmed the validity of the proposed solution.

4 Results

4.1 Selection and Review of Articles

The types of articles identified include:

- **Technical:** A tool, algorithm, or specific method designed to solve a problem, with detailed implementation aspects.
- **Strategy:** A general approach to addressing a problem, defining an action plan with key steps.
- **Procedure:** A series of sequential steps to accomplish a task, presented with clear instructions.
- **Methodology:** A set of methods and principles that establish a systematic framework for projects or research.
- **Case Description:** A real-world example of an application, including the context, problem, solution, and results.
- **Recommendation:** Guidance for decision-making in future developments.

The review repeatedly identified the use of terminology related to structure, interaction, and behavior. Table 1 highlights the articles that reference these aspects.

4.2 Content Analysis

The content analysis conducted on each of the articles identified a total of 22 categories, which were grouped into six main themes characterized by recurring elements present in the analyzed texts. These themes were classified into three main groups: structure, interaction, and system behavior (see Fig. 1).

Table 1. Collected and analyzed articles

#	Article	Year	Type	Domain	Structure	Interaction	Behavior
p1	[23]	2023	Case description	Education	x		x
p2	[21]	2024	Case description	Networking		x	
p3	[24]	2023	Case description	Epidemiology	x		
p4	[18]	2024	Strategy	Multidomain	x	x	
p5	[30]	2020	Technique	Data Visualization	x		x
p6	[19]	2024	Recommendation	Geoinformatics	x		x
p7	[17]	2019	Technique	Smart Cities		x	
p8	[26]	2017	Procedure	Manufacturing		x	
p9	[16]	2022	Methodology	Public Safety	x		x
p10	[22]	2023	Methodology	Local Governments		x	
p11	[13]	2019	Methodology	Smart Cities		x	
p12	[32]	2020	Methodology	Public Health	x		
p13	[5]	2020	Methodology	Public Health		x	x
p14	[6]	2021	Case description	Land Management	x		

4.3 Identification of Key Elements and Operational Dynamics

Figure 2 illustrates the proposed elements for each group, derived from the content analysis. The details of each component are presented in Table 2.

This initial classification enabled the identification of predominant elements that, when combined, allow for the generation of views (see Fig. 2):

1. **Structure:** Refers to the foundation that facilitates user actions and ensures system responses are coherent. This block encompasses: Let E be the set of all elements that comprise the structure of a dashboard:

$$E = \{e_1, e_2, e_3, \ldots, e_{|E|}\} \tag{1}$$

   Examples of e_k include: the main map area, navigation panel, filter panel, legends and layers, header, and footer.
2. **Interaction:** The set of actions users can perform on the components of a system. It determines the level of usability and accessibility, directly influencing how users achieve their goals within the system. Let I be the set of all possible interactions defined on structural elements:

$$I = \{i_1, i_2, i_3, \ldots, i_{|I|}\} \tag{2}$$

   Examples of i_k include: zoom, map panning, clicking on a marker, selecting a filter, performing a geographic search, toggling layers on or off, etc.
3. **Behavior:** The internal responses and transformations that occur within a system as a result of actions or events. It includes state changes and the reactions triggered by stimuli, ensuring coherence in how the system processes and responds to different situations, impacting efficiency and user satisfaction. Let B be the set of reactions or behaviors the system can exhibit in response to interactions:

$$B = \{b_1, b_2, b_3, \ldots, b_{|B|}\} \tag{3}$$

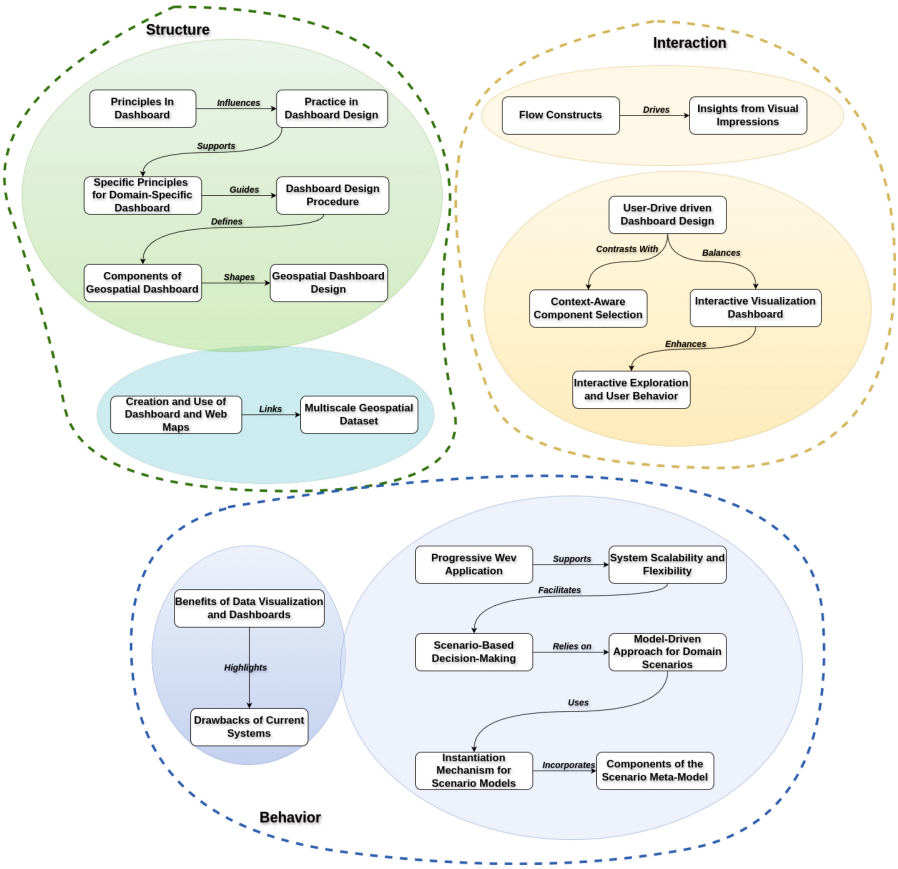


Fig. 1. Identified categories

Examples of b_k include: real-time data updates, highlighting an area, displaying a tooltip, transition animations, error notifications, loading indicators, etc.

4. **Views:** Configurations or tangible representations of the system displayed to the user at a given moment. These integrate the structure layout, interaction options, and active behaviors, reflecting the current state of the interface. Views enable users to access and manipulate information concretely by synthesizing the components defined in the structure, available actions, and system responses into a visual and functional experience. Let V represent the set of all possible views (or screens) resulting from combining structure, interaction, and behavior at a given moment:

$$V = \{v_1, v_2, v_3, \ldots, v_{|V|}\} \tag{4}$$

Each $v \in V$ represents a configuration/instance of the screen as seen by the user (with certain elements visible, specific filters active, etc.).

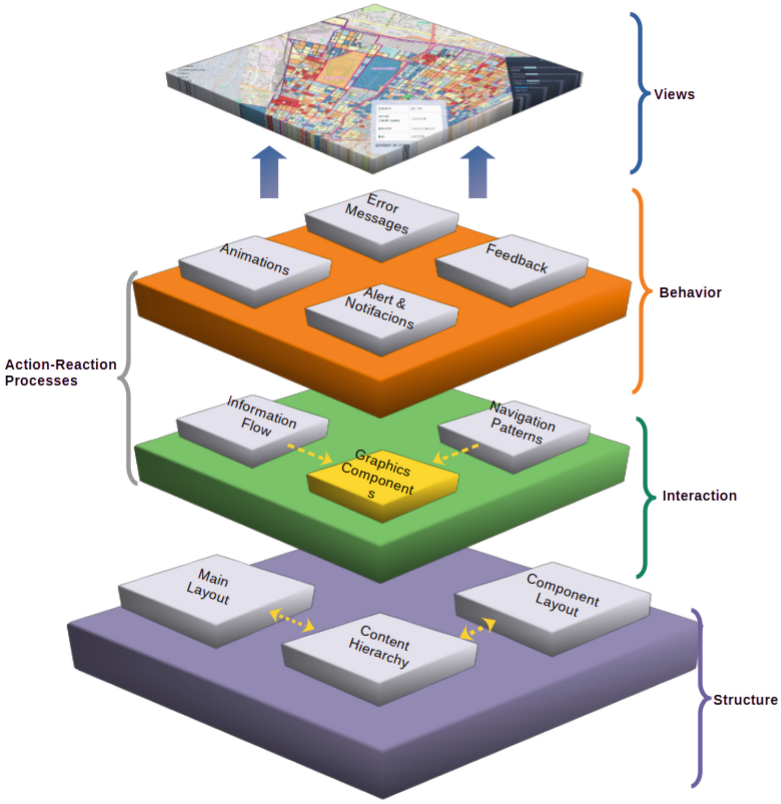


Fig. 2. Relevant Factors

It is worth mentioning that the identified elements E, I, C are consistent with the types of elements focused on in usability and human-computer interaction testing described in [28], and can also be partially identified in the proposal outlined in [4].

View Generation. The dynamics resulting from user actions and the corresponding system responses are referred to as the Action-Reaction Process, enabling fluid and bidirectional communication between the system and its users, thereby influencing the experience and effectiveness of the interaction.

We can assume that each view v arises from a combination (or multiple combinations) of Structure, Interaction, and Behavior elements. To model this, we define a generic function:

$$\Phi : \mathcal{P}(E) \times \mathcal{P}(I) \times \mathcal{P}(B) \rightarrow V \tag{5}$$

Table 2. Relevant Elements Associated with Logical Blocks

Parts	Section	Description
Structure	Main Layout	Spatial organization of the map and panels, prioritizing accessibility and clarity. Key elements, such as the map, occupy the center for immediate visual focus.
	Content Hierarchy	Organization of data by importance; key indicators first, followed by complementary charts, and additional details in secondary areas.
	Component Distribution	Strategic positioning of interactive and visual elements to maximize functionality and reduce cognitive load; the map serves as the core, with filters and data surrounding it.
Interaction	Information Flow	Logical sequence connecting map data with charts and filters, enabling intuitive analysis from general views to specific details in real time.
	Graphic Components	Interactive filters, combo boxes, search fields, sliders, and buttons designed to manipulate, explore, and customize data presented on the map and associated charts.
	Navigation Patterns	Smooth transitions between components, prioritizing minimal clicks, contextual navigation from the map to details, and intuitive backtracking for exploring multiple levels of information.
Behavior	Feedback	Immediate and clear system responses to user actions, such as data updates, highlighted areas on the map, or visual confirmations.
	Alerts and Notifications	Visual or auditory indicators highlighting relevant events, such as critical updates, data errors, map changes, or important warnings.
	Animations	Smooth transitions between views, graphical effects during interactions with maps or filters, and dynamic highlighting to emphasize changes or interactions.
	Error Messages	Clear and precise indications explaining problems, suggesting solutions, or guiding users to correct improper actions without causing frustration.

where:

- $\mathcal{P}(X)$ denotes the power set of X, i.e., all possible subsets of X.
- $\Phi(E', I', B')$ incorporates structure by selecting and organizing elements, assigns interactions by associating actions, and configures behavior by defining responses. This process produces a view or set of views by selecting:
  - $E' \subseteq E$: the active/visible structural elements.
  - $I' \subseteq I$: the enabled interactions.
  - $B' \subseteq B$: the available behaviors or responses.

Constraints. In practice, the combinations (E', I', B'), originate from requirements stage, represents the valid and meaningful structural elements enabling coherent interactions and specific behaviors. Thus, certain behaviors only apply when appropriate structural elements are present, and interactions depend on specific designed behaviors.

Thus, it is possible to establish certain constraints:

- **Dependency between Structure and Interaction**

$$R_{E,I} \subseteq E \times I$$

  which indicates, for example, that the interaction i_j only applies if the structural element e_i is present in the view.
- **Dependency between Interaction and Behavior**

$$R_{I,B} \subseteq I \times B$$

  to describe that a certain behavior b_k is the specific response to the interaction i_j.
- **Global Dependency** By combining everything, we can define a relationship

$$R \subseteq E \times I \times B$$

  which summarizes the compatibility between a structural element $e \in E$, an interaction $i \in I$, and a behavior $b \in B$.

In such cases, the function Φ could be restricted to:

$$\Phi : \{(E', I', B') \mid (E', I', B') \text{ satisfy } R\} \rightarrow V$$

to ensure that only valid views are possible.

Algorithm. Algorithm 1 describes how the components of Structure, Interaction, and Behavior are defined, validating possible combinations according to the constraints and generating the views. In phases 1, 2, and 3, the elements of the Structure layer (E), Interaction (I), and Behavior (B) are defined. In phase 4, the constraints between the elements of E, I, and B are established. In phase 5, the views are generated by iterating through all subsets of E, I, and B. For each combination (E', I', B'), it is verified whether the combination is valid according to R. The function Φ produces a view v. Finally, the set of all compatible views is returned.

4.4 Preliminary Implementation of the Proposal in a Pilot

The key identified elements were applied in the development of a pilot system designed as a georeferenced analytics dashboard for multiple deprivation indicators (MDI). This concept, which encompasses various dimensions of analysis such as material and social deprivation, is fundamental for assessing poverty and health conditions in specific areas and supporting data-driven public policy decisions [1].

Algorithm 1. View Generation in a Dashboard with Maps

Require: Sets:
$E = \{e_1, \ldots, e_m\}$ *(Structural Elements)*
$I = \{i_1, \ldots, i_n\}$ *(Interaction Elements)*
$B = \{b_1, \ldots, b_p\}$ *(Behavioral Elements)*
Constraint Relation: $R \subseteq E \times I \times B$ *(Compatibility)*
View generation function: $\Phi : \mathcal{P}(E) \times \mathcal{P}(I) \times \mathcal{P}(B) \to V$
Ensure: Set of Views: $V_{\text{final}} \subseteq V$

Phase 1: Structure Definition
Define the set E of structural elements
Phase 2: Interaction Definition
Define the set I of possible interactions
Phase 3: Behavior Definition
Define the set B of behaviors
Phase 4: Constraint Configuration
Determine the relation $R \subseteq E \times I \times B$ that indicates
  which combinations of (e, i, b) are valid.
Phase 5: View Generation
Initialize $V_{\text{final}} \leftarrow \varnothing$
Phase 5.1: Incremental Subset Selection
for all $E' \subseteq E$ **do**
  (a) Activated Structure: E'
  for all $I' \subseteq I$ **do**
    (b) Available Interactions: I'
    for all $B' \subseteq B$ **do**
      (c) Active Behaviors: B'
      Verify if for all (e, i, b) in $E' \times I' \times B'$,
      it holds that $(e, i, b) \in R$ **when applicable**.
      if the combination (E', I', B') **is compatible with** R **then**
        Generate the view $v \leftarrow \Phi(E', I', B')$
        $V_{\text{final}} \leftarrow V_{\text{final}} \cup \{v\}$
      end if
    end for
  end for
end for
Phase 6: Result
return V_{final} *(Set of all valid views)*

Definition of the Sets. For a complex dashboard such as the pilot, it is possible to define the elements of the structure vector E outlined in (1):

e_1: base map corresponding to the main area where the map is displayed.
e_2: side navigation panel to move between dashboard sections.
e_3: left filter panel related to geographic distribution.
e_4: section explaining the map's symbology and layer management.
e_5: right-hand side panel displaying detailed information about the area selected on the map.
e_6: blocks containing key indicators (numbers and global metrics).

The set of elements for the interaction vector I outlined in (2):

i_1: applying options in the filter panel (e_3).
i_2: zooming in or out on the map (e_1) using buttons or mouse scroll.
i_3: panning and dragging the map (e_1).
i_4: clicking on a point on the map (e_1).
i_5: selecting an area (polygon) on the map (e_1) for more details (e_5).
i_6: displaying layers with dimensions of the indicator in question (MDI).
i_7: hiding layers with dimensions of the indicator in question (MDI).
i_8: using a text field to search for a location and center the map on it.

Finally, for the behavior vector B described in (3), the following elements are defined:

b_1: reloading information on the map (e_1) and/or panels (e_2 and e_5) when changing filters (e_3) or zooming (i_2).
b_2: changing the marker's color, size, etc., when hovering or clicking on it.
b_3: smooth animation/transition when zooming (i_2).
b_4: displaying a brief information box when hovering over an element (tooltip).
b_5: showing a popup window (with additional data) when clicking on an area or marker (PopUp).
b_6: overlaying new data layers.
b_7: error messages in case of an issue (connection failure, invalid filters).
b_8: effects when showing/hiding panels (e_2, e_3, e_5) (smooth slide, *fade*, etc.).

View Generation Function: Using the view generation function (5), which, given a combination of structural elements (E'), interactions (I'), and behaviors (B'), generates (or describes) a view $v \in V$.

View 1: "General View"

- **Active Structure (E'_1):**

$$E'_1 = \{e_1, e_2, e_3, e_4, e_6\}.$$

- **Available Interaction (I'_1):**

$$I'_1 = \{i_1, i_2, i_3, i_6, i_7, i_8\}.$$

- **Configured Behavior (B'_1):**

$$B'_1 = \{b_1, b_3, b_7, b_8\}.$$

The resulting view can be expressed as:

$$v_1 = \Phi(E'_1, I'_1, B'_1).$$

View 2: "Marker Detail"

- **Active Structure (E'_2):**

$$E'_2 = \{e_1, e_2, e_3, e_4, e_6, e_5\}.$$

- **Available Interaction (I'_2):**

$$I'_2 = \{i_4, i_2, i_3, i_6, i_7, i_8\}.$$

- **Configured Behavior (B'_2):**

$$B'_2 = \{b_2, b_4, b_5, b_1, b_3\}.$$

The resulting view can be expressed as:

$$v_2 = \Phi(E'_2, I'_2, B'_2).$$

Part of view v_1 can be reviewed in Fig. 3.

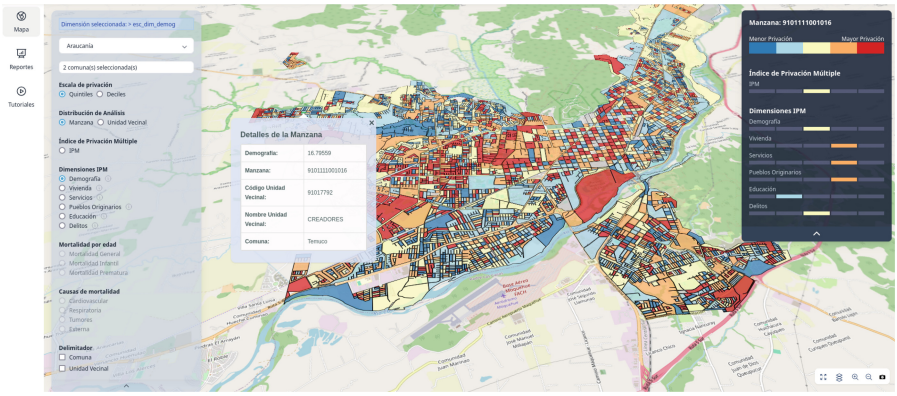


Fig. 3. General view created from View 1

Relationships and Constraints: In the case of the pilot, not all combinations of (E', I', B') make sense. For instance, the following constraints can be detailed:

- **Structural Constraint:** Interaction i_4 cannot occur if e_1 with markers does not exist:

$$\text{if } e_1 \notin E', \text{ then } i_4 \notin I'.$$

- **Behavioral Constraint:** b_2 depends on the existence of interaction i_4 on a marker. If this interaction does not exist, there is no reason to activate b_2.

- **Layer Dependency:** i_6 and i_7 only make sense if e_4 exists to display the layers.
- **Dependency Relationships:** Dependency relationships are thus R_E, $R_{I,E}$, $R_{I,B}$, or:

$$R \subseteq E \times I \times B.$$

In this way, we ensure that Φ only generates valid views.

5 Conclusions

This research analyzed the key factors for the design and implementation of a georeferenced data analysis dashboard, as well as the optimal sequencing of these elements. Through the collection and analysis of scientific literature, three fundamental blocks structuring the critical factors were identified: structure, interaction, and behavior of the system components. This approach is complemented by an implementation procedure aimed at providing both manual and automated support, integrating the potential use of artificial intelligence agents. In this context, the view generation function is susceptible to automation through AI agents capable of autonomously projecting a dashboard system. Additionally, it can be used in testing phases to detect unforeseen interaction combinations or potential optimization paths in interaction sequences, among other HCI aspects.

It is important to note that this work is in a preliminary stage, and definitive results will be obtained after the operational deployment of the analysis dashboard, where this proposal has been implemented.

We identified two main limitations in this study. First, at the methodological level, it is essential to validate the proposal in other similar systems to strengthen its robustness and applicability. Second, there is a potential bias stemming from the predominant focus on the macrostructure of the system, relegating the user perspective, which was considered primarily through the literature review. This latter aspect will be addressed subsequently through standardized usability testing in the pilot system, allowing the results to be contrasted with our initial proposal.

Acknowledgments. Jaime Díaz-Arancibia is supported by Grant ANID, Chile, Fondecyt de Iniciación en Investigación, Project N° 11230141

References

1. Alkire, S., Foster, J.: Understandings and misunderstandings of multidimensional poverty measurement. J. Econ. Inequality **9**(2), 289–314 (2011). https://doi.org/10.1007/s10888-011-9181-4
2. Almasi, S., Bahaadinbeigy, K., Ahmadi, H., Sohrabei, S., Rabiei, R.: Usability evaluation of dashboards: a systematic literature review of tools. BioMed Res. Int. **2023**(1), 9990933 (2023). https://doi.org/10.1155/2023/9990933

3. Arksey, H., O'Malley, L.: Scoping studies: towards a methodological framework. Int. J. Soc. Res. Methodol. **8**(1), 19–32 (2005). https://doi.org/10.1080/1364557032000119616
4. Bach, B., et al.: Dashboard design patterns. IEEE Trans. Vis. Comput. Graph. **29**(1), 342–352 (2023). https://doi.org/10.1109/TVCG.2022.3209448, https://ieeexplore.ieee.org/document/9903550, conference Name: IEEE Transactions on Visualization and Computer Graphics
5. Bobashev, G., Segovia-Dominguez, I., Gel, Y.R., Rineer, J., Rhea, S., Sui, H.: Geospatial forecasting of COVID-19 spread and risk of reaching hospital capacity. SIGSPATIAL Special **12**(2), 25–32 (2020). https://doi.org/10.1145/3431843.3431847
6. Briggs, M.: SatDash: an interactive dashboard for assessing land damage in Nigeria and Mali. In: Proceedings of the 4th ACM SIGCAS Conference on Computing and Sustainable Societies, pp. 100–114. COMPASS '21, Association for Computing Machinery, New York, NY, USA (2021). https://doi.org/10.1145/3460112.3471949
7. Bylinskii, Z., et al.: Learning visual importance for graphic designs and data visualizations. In: Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology, pp. 57–69. UIST '17, Association for Computing Machinery, New York, NY, USA (2017). https://doi.org/10.1145/3126594.3126653
8. Contreras-Figueroa, V., Montané-Jiménez, L.G., Cepero, T., Benítez-Guerrero, E., Mezura-Godoy, C.: Information visualization in adaptable dashboards for smart cities: a systematic review. In: 2021 9th International Conference in Software Engineering Research and Innovation (CONISOFT), pp. 34–43 (2021). https://doi.org/10.1109/CONISOFT52520.2021.00017, https://ieeexplore.ieee.org/document/9653197
9. Creswell, J.W., Creswell, J.D.: Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. SAGE Publications Inc, Thousand Oaks, California (2023)
10. Ding, W., Lin, X., Zarro, M.: Information architecture and UX design: definitions and scope. In: Ding, W., Lin, X., Zarro, M. (eds.) Information Architecture and UX Design: The Integration of Information Spaces, pp. 1–9. Springer Nature Switzerland, Cham (2025). https://doi.org/10.1007/978-3-031-72138-0_1
11. Few, S.: Information Dashboard Design: The Effective Visual Communication of Data. Inc, O'Reilly Media (2006)
12. ISO/IEC/IEEE: ISO 9241-210:2019(en), Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems (2019). https://www.iso.org/obp/ui/en/#iso:std:iso:9241:-210:ed-2:v1:en
13. Jing, C., Du, M., Li, S., Liu, S.: Geospatial dashboards for monitoring smart city performance. Sustainability **11**(20), 5648 (2019). https://doi.org/10.3390/su11205648, https://www.mdpi.com/2071-1050/11/20/5648, number: 20 Publisher: Multidisciplinary Digital Publishing Institute
14. Kirk, A.: Data Visualisation: A Handbook for Data Driven Design. SAGE Publications Ltd (2019)
15. Lance, A. Waller, C.A.G.: Applied Spatial Statistics for Public Health Data. Willey & Sons (2004)
16. Liu, Z.G., Li, X.Y., Zhu, X.H.: scenario modeling for government big data governance decision-making: Chinese experience with public safety services. Inf. Manage. **59**(3), 103622 (2022). https://doi.org/10.1016/j.im.2022.103622, https://www.sciencedirect.com/science/article/pii/S0378720622000349

17. Mena, M., Corral, A., Iribarne, L., Criado, J.: A progressive web application based on microservices combining geospatial data and the internet of things. IEEE Access **7**, 104577–104590 (2019). https://doi.org/10.1109/ACCESS.2019.2932196, https://ieeexplore.ieee.org/document/8782470, conference Name: IEEE Access
18. Nasir, A.H.M., Surin, E.S.M., Ahmad, M.N.: Towards unified information dashboard design - common principles, practices, and challenges. Procedia Comput. Sci. **234**, 1586–1592 (2024). https://doi.org/10.1016/j.procs.2024.03.161, https://www.sciencedirect.com/science/article/pii/S1877050924005209
19. Netek, R.: Map-oriented dashboards online-from minor method to GIScience trend. In: Yadava, R.N., Ujang, M.U. (eds.) Advances in Geoinformatics Technologies : Facilities and Utilities Optimization and Management for Smart City Applications, pp. 33–57. Springer Nature Switzerland, Cham (2024). https://doi.org/10.1007/978-3-031-50848-6_3
20. Paz, F., Pow-Sang, J.A.: Usability evaluation methods for software development: a systematic mapping review. In: 2015 8th International Conference on Advanced Software Engineering & Its Applications (ASEA), pp. 1–4 (2015). https://doi.org/10.1109/ASEA.2015.8, https://ieeexplore.ieee.org/document/7433059
21. Saeed, K.A., Wu, D., Xu, D.J.: Effect of designer- versus user-driven network-monitoring dashboard design on user flow experience and performance: the role of augmented virtuality. Inf. Manage. **61**(3), 103926 (2024). https://doi.org/10.1016/j.im.2024.103926, https://www.sciencedirect.com/science/article/pii/S0378720624000089
22. Stupen, R., Ryzhok, Z., Stupen, N., Stupen, O.: Methodology for creating geoportals of local self-government bodies using geoinformation systems. In: 2023 IEEE 18th International Conference on Computer Science and Information Technologies (CSIT), pp. 1–4 (2023). https://doi.org/10.1109/CSIT61576.2023.10324151, https://ieeexplore.ieee.org/document/10324151, iSSN: 2766-3639
23. Sudharson, K., Sudhire Rahul, K., Sai Krishna, R.G.: Interactive dashboard for education abroad with geospatial information and predictive analysis. In: 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 1067–1072 (2023). https://doi.org/10.1109/ICICCS56967.2023.10142437, https://ieeexplore.ieee.org/document/10142437, iSSN: 2768-5330
24. Usman, M., Zhou, H., Moon, S., Zhang, X., Faloutsos, P., Kapadia, M.: A multiscale geospatial dataset and an interactive visualization dashboard for computational epidemiology and open scientific research. IEEE Comput. Graph. Appl. **43**(1), 39–52 (2023). https://doi.org/10.1109/MCG.2022.3230444, https://ieeexplore.ieee.org/document/9992069, conference Name: IEEE Computer Graphics and Applications
25. Vermeeren, A.P.O.S., Law, E.L.C., Roto, V., Obrist, M., Hoonhout, J., Väänänen-Vainio-Mattila, K.: User experience evaluation methods: current state and development needs. In: Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, pp. 521–530. NordiCHI '10, Association for Computing Machinery, New York, NY, USA (2010). https://doi.org/10.1145/1868914.1868973
26. Vilarinho, S., Lopes, I., Sousa, S.: Design procedure to develop dashboards aimed at improving the performance of productive equipment and processes. Procedia Manuf. **11**, 1634–1641 (2017). https://doi.org/10.1016/j.promfg.2017.07.314, https://www.sciencedirect.com/science/article/pii/S235197891730522X
27. Vázquez-Ingelmo, A., Garcia-Peñalvo, F.J., Therón, R.: Information dashboards and tailoring capabilities - a systematic literature review. IEEE Access **7**,

109673–109688 (2019). https://doi.org/10.1109/ACCESS.2019.2933472, https://ieeexplore.ieee.org/document/8789402
28. Washizaki, H.: Guide to the software engineering body of knowledge (SWEBOK Guide) version 4.0. Tech. rep., IEEE Computer Society (2024). www.swebok.org
29. Yigitbasioglu, O.M., Velcu, O.: A review of dashboards in performance management: Implications for design and research. Int. J. Account. Inf. Syst. **13**(1), 41–59 (2012). https://doi.org/10.1016/j.accinf.2011.08.002, https://www.sciencedirect.com/science/article/pii/S1467089511000443
30. Yu, J., Chowdhury, K., Sarwat, M.: Tabula in action: a sampling middleware for interactive geospatial visualization dashboards. Proc. VLDB Endow. **13**(12), 2925–2928 (2020). https://doi.org/10.14778/3415478.3415510
31. Zimmerman, J., Forlizzi, J., Evenson, S.: Research through design as a method for interaction design research in HCI. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 493–502. CHI '07, Association for Computing Machinery, New York, NY, USA (2007). https://doi.org/10.1145/1240624.1240704
32. Çay, D., Nagel, T., Yantaç, A.E.: Understanding user experience of COVID-19 maps through remote elicitation interviews. In: 2020 IEEE Workshop on Evaluation and Beyond - Methodological Approaches to Visualization (BELIV), pp. 65–73 (2020). https://doi.org/10.1109/BELIV51497.2020.00015, https://ieeexplore.ieee.org/document/9307749

A Serious Game Approach for Teaching Requirements Engineering: User Experience Evaluation

Tomás Castillo[1], Benjamín Chamorro[1], Ignacio Catalán[1], Pablo Schwarzenberg[1(✉)], Luis Rojas[3], and Fei Luo[2]

[1] Facultad de Ingeniería, Universidad Andrés Bello, Santiago, Chile
pablo.schwarzenberg@unab.cl
[2] School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China
luof@ecust.edu.cn
[3] Facultad de Ingeniería, Universidad San Sebastián, Santiago, Chile
luis.rojasp@uss.cl

Abstract. Requirement engineering is one of the key processes in software development, being misunderstandings between the project team and relevant stakeholders the source of some of the most critical problems in software development projects. To teach requirement engineering is challenging, because it is a combination of communicational and analytical skills, and it is the reason because most courses use various approaches to facilitate learning, from theory-oriented lectures to active learning using role playing, simulations or project-based learning. In online education, the challenge of effectively teaching requirement engineering increases, because most educational programs are mainly delivered asynchronously with few synchronic opportunities for engagement in collaborative active learning. This work proposes the use of a serious game to teach the fundamentals of requirement engineering to overcome these limitations. The game is a card game in a simulated context where the interactions of the students with the game are recorded. A usability evaluation was conducted to assess the game experience, being the game evaluated positively in all proposed dimensions, standing out in the dimensions of Ease of Learning, Clarity of Objectives, and Educational Concepts. The game's strong dependence on luck was qualitatively highlighted, but also the ease of use and a positive user experience, proving to be effective in engaging students' interest.

Keywords: Requirements Engineering · Serious Games · Game-based learning

1 Introduction

Requirements Engineering (RE) can be defined as the set of processes related to requirements, from discovery, development, analysis and prioritization to imple-

T. Castillo, B. Chamorro, and I. Catalán—These authors contributed equally to this work.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 20–36, 2025.
https://doi.org/10.1007/978-3-031-93539-8_2

mentation, verification and validation. Their purpose is the management of the requirements that define the system at different levels of abstraction [6]. RE is an area of Software Engineering that is crucial to the success of a project, because of that it is usually given greater emphasis when teaching Software Engineering. However, its teaching/learning process is challenging, because combine computer science and social science concepts [18]. If the teaching it is done via traditional methods, the disconnection between theory and experience makes it difficult to internalize its concepts and the ability to apply them.

Multiple approaches had been proposed to teaching requirement engineering [5], from project-based learning (PBL) in real contexts [21] to role-playing [15]. Participating in real projects gives students opportunities to develop the communication skills needed for RE, but it is difficult to set up [18], especially in online education. Because of that, role-playing emerges as an effective solution for resource constrained contexts [18].

Serious games provide opportunities for exploration and role-playing to students, as a preparation for real life roles [1], and can be a solution to implement role-playing to improve RE teaching.

A Serious Game is a game whose primary purpose is educational and not entertainment [1], although it must be fun and motivating to be effective. Serious games allow the communication of concepts and facts of many subjects [1].

Serious Games can be very effective for learning [9,14,26,27], having several advantages over traditional learning methods. They allow for hands-on learning, giving the user a safe space to experiment, learn from mistakes, and practice skills; they can be motivating, the playful elements make users more willing to participate and learn; and they can be highly customizable to suit the user's pace and needs.

2 Related Works

Serious games had a long history of use in education as tools for learning ([1,4, 20]) and assessment ([3,19,22]). In the context of requirement engineering (RE) education, a set of 12 previous works were studied. A summary of the studies analyzed can be seen in Table 1. The studies cover the years 2008 to 2023. The relatively short 16-year period in which they fall accounts for the relatively recent emergence of interest in the application of serious games to requirement engineering education. The games analyzed can be separated into two categories:

- **Games that teach RE skills through simulation.** This category includes the games TREG [25], *Untitled Game* [10], Biyubi [7], *Untitled Game* [11], Requengin [8], and BARA [17].
- **Games that didactically assess previously acquired general skills.** This category includes the games Quantum Software [13], RE-O-Poly [23], Guess What we Want [2], ReGo [2], Q & A & Q [2], Communication and Traceability Game [12], The Requirement Island [24] and UbiRE [16].

In the first category, two prominent trends were identified in the existing literature to date: (1) the skill most often attempted to be taught is elicitation,

that related to practices on how to do an effective initial stakeholder requirements gathering; and (2) the style most often used is role-playing, a style of play that puts the player in the shoes of a character who must interact with NPCs (Non-Playable Characters).

The first observation indicates a common convention that the initial phases of Requirements Engineering are the most important to ensure the future of the project; the student must be able to efficiently identify the necessary requirements for a project, which will then be the pillars of the project. In this type of game, the player must traverse a virtual space populated with different characters with defined needs and desires. He must interact with the different characters to discover their needs and desires, and discern which ones deserve to be considered as requirements for the project. In some games, the player must then define and refine these requirements into formal requirements.

The second observation makes sense when it is understood that the ability to take good requirements goes hand in hand with the need to communicate effectively with external people and stakeholders. The role-playing game style provides the instance to simulate the necessary social relationships in a safe environment for the learner. In most games in this category, the player must begin by casually interacting with the characters and then conduct interviews to obtain more accurate information.

It is noteworthy the fact that during the last 5 years there has been a trend towards role-playing games, probably because of the similarity that can be achieved with a real situation, which has proven to be more efficient than other styles of games as an educational tool.

Role-playing games such as TREG [25], Biyubi [7], Requengin [7] and BARA [17] are very similar in the way they approach elicitation activities; they present the player with a series of characters with a level of complexity necessary to provide variety in their interaction, each with different personal interests, ideas, and needs to be taken into account as requirements for the system the game asks the player to develop, in addition to the mechanics necessary to obtain that information, such as conversations, interviews, and visualization of character statistics.

In all the above cases, the realism provided by this variety of characters helps the player to develop critical analysis skills to make decisions in a realistic environment, but without real risk. This approach has proven to be efficient for learning and for preparing the student for a real work environment.

However, a common factor of role-playing games is that, in the final phases of the created requirements' analysis, no game performs an automatic evaluation of the player's performance but instead conclude with an external evaluation of the game, performed by the teacher in charge of the student who is playing, with data exported from the game that summarizes the student's actions and decisions. This makes the effectiveness of the game dependent on a course context, with students and a teacher.

As for the second category, games such as RE-O-Poly [23], ReGo [2] and The Requirements Island [24] are games that, instead of attempting to teach concepts,

seek to evaluate knowledge previously acquired by traditional methods. They are games that by themselves are incomplete and function only as a complement to previous teaching, but unlike the games in the first category, they are capable of automatically evaluating the user's performance.

From this analysis, the obstacle of creating a game that can teach the concepts of ER effectively without becoming cumbersome was identified; this was evidenced in several examples, highlighting the case of ReGo [2], in which students presented divided opinions: on the one hand, they are a good training tool and a more didactic way to "digest" the content, and on the other hand, a balance must be found so that it does not become a distraction and the focus is not lost, finding a more didactic way to deliver knowledge.

Table 1. Summary of Contemplated Previous Studies.

Reference	Title	Genre	Activity	Participants
Knauss et al., 2008 [13]	Software Quantums	Puzzle	General RE practices	Not specified
Smith, 2008 [23]	RE-O-Poly	Board game	General RE practices	49
Alexander, 2008 [2]	Guess What We Want	Strategy	Management	Not specified
Alexander, 2008 [2]	ReGo	Board game	General RE practices	Not specified
Alexander, 2008 [2]	Q & A & Q	Logic	Elicitation	Not specified
Vega et al., 2009 [25]	TREG	Role-playing game	Elicitation	Not specified
Jaramillo, 2010 [12]	Communication and Traceability Game	Card game	Elicitation	37
Thiry, 2010 [24]	The Requirement Island	Puzzle	General RE practices	+40
Hainey et al., 2011 [10]	Untitled	Role-playing game	Elicitation, Analysis	174
Lima et al., 2012 [16]	UbiRE	Task management	General RE practices	5
García et al., 2019 [7]	Biyubi	Role-playing game	Elicitation	94
Ibrahim et al., 2019 [11]	Untitled	Role-playing game	Elicitation, Analysis	10
García et al., 2020 [8]	Requengin	Role-playing game	General RE practices	80
Liu et al., 2023 [17]	BARA	Role-playing game	Elicitation	60

Therefore, we propose a gamified approach to requirement engineering education, with a card game that allows Computer Engineering students to apply the requirement engineering skills acquired during a Software Engineering course in a simulated context. Because a motivating gameplay it is crucial for game success and usability it is indispensable for game usefulness as a teaching tool, we propose the following research questions to guide this study:

RQ1: How motivating is the developed game in maintaining student engagement throughout the learning process?

RQ2: How effective is the game in achieving its educational objectives across different dimensions, including clarity, usability, and challenge?

3 Proposal

The Serious Game proposed in this study to evaluate Requirements Engineering concepts is called ***Req Duel***, it is a card game inspired by the gameplay of video games such as *Hearthstone* and *Yu-Gi-Oh! Master Duel*.

Req Duel presents the player with everyday problems in the context of a Metro transportation system. The player has a deck of predetermined cards, from which they must draw each turn to add to their hand, and they must manage both the resources in their hand and financial resources to solve the problems presented to them each turn in the form of enemy cards called Stakeholders. The player's objective is to solve the problems presented to them during 10 consecutive turns (represented as days), preventing any of the Stakeholders from losing their patience.

Fig. 1. Game Screenshot ***Req Duel***

In Fig. 1 we can see the elements that make up the game:

(**1**) At the top of the board are the Stakeholder cards. These cards contain the problems that the player must solve and are represented by the color red. They consist of a representative image, a value in the center representing their patience level, and their name. Patience drops by 1 at the start of each turn; if the patience of any of the Stakeholders reaches zero, the player loses. A new Stakeholder card will appear at the beginning of each turn to be added to those already present. By hovering the mouse over a Stakeholder card, a description of the problem to be solved can be displayed, as shown in the Stakeholder card on the left.

(**2**) At the bottom of the board is the player's hand. These are the cards that the player must use to solve the problems presented. There are two types of cards:

- In green are represented the Solution cards, these are the main cards of the hand. During the turn, the player can select a Solution card, followed by a Stakeholder card that he/she believes can be solved with the chosen Solution card.

- In purple are represented the Mitigation cards, these are cards that provide indirect help for the player, with effects such as temporarily increasing the budget, drawing extra cards, etc.

All cards in the player's hand, both Solutions and Mitigations, have a cost to use, represented by the value in the center of the card, this cost is subtracted from the budget the player has available during the turn. If the cost of a card is greater than the remaining budget, the player will not be able to play it during that turn. Like the Stakeholder cards, if the player hovers the mouse over any of these cards, the effect of the card will be shown.

(**3**) The main elements of the user interface are located on the right side of the dashboard:
- Calendar: represents the player's current turn. Each turn symbolizes a Day of work. When the player manages to finish Day 10, they win the game.
- Budget: This is the resource with which the player can play his cards. The first value represents the available budget, the second represents the total budget. At the start of each turn, the total budget increases by one, and the available budget is restored to equal the total. If the player spends all the available budget during a turn, he will not be able to play any more cards and will be forced to pass to the next turn by pressing the Pass Day button.
- Pass Day button: the player can press this button at any time to end the current turn and start the next one. At the beginning of each turn the budget will be reset as explained in the previous item; a new Stakeholder card will come into play, and two new cards will be added to the player's hand.

(**4**) This is the player's deck, which shows how many cards are available. The game starts with 24 Solution cards and 14 Mitigation cards for a total of 38 cards that are randomly shuffled into this deck.

(**5**) This space is the discard area, where all the cards that the player has used during the game are placed.

(**6**) A small Help icon that the player can consult at any time by clicking on it. A menu will open that briefly explains each of the elements of the game.

4 Method

The study's design is quasi-experimental, observational, and single group, given that it was tested in a controlled academic setting. All participants belonged to a course at a Chilean university and tested the game during an in-person session in the classroom. The test was conducted with each student using their own computer, on which the game was installed. After completing the game session, the participants answered the evaluation forms individually. There were no divisions into control groups or subsequent follow-ups; the activity was designed as a single interactive experience with immediate data collection; therefore, there is a single sample with no variations. Students were informed of the study and invited to participate on the same day, and the activity was conducted in-person in a classroom. Data collection was completed at the end of the session, after the game test. The students were mostly male, had an average age of 22 years had a basic Requirements Engineering knowledge.

A total of 26 students were present during the session, testing the game and answering the assessment surveys. Two questionnaires were applied to the students, and were created specifically for this study, and included questions with Likert scale responses from 1 to 5, as well as open-ended questions. Participants responded in a controlled environment, under supervision of the investigators, to solve questions regarding the surveys. Blinding was not used, as the aim of the study was to collect students' direct perception of the game, and the awareness of the study participants had no influence on their responses.

4.1 Form 1: REQ DUEL Usability Questionnaire

This questionnaire evaluated aspects of usefulness, usability, learning, and overall satisfaction. The questions were classified into four dimensions for analysis:

- **Utility:** Responds to the game's ability to improve aspects of the player's personal life, such as productivity or satisfaction of needs.
- **Ease of Use:** Responds to the user experience when playing, considering the level of complexity of use, consistency and adaptability of the application.
- **Ease of Learning:** Responds to the ease and speed with which users can understand and master the use of the application.
- **Satisfaction:** Responds to the user's general feeling towards the application, the positivity of the experience and the engagement.

The *FORM1: REQ DUEL Usability Questionnaire* contains 31 Likert-scale response questions from 1 to 5, where 1 means "Strongly Disagree" and 5 means "Strongly Agree", separated into 4 dimensions to analyze. It also contains two open-ended questions highlighting negative and positive aspects of the game respectively.

4.2 Form 2: Game Evaluation Survey

This questionnaire evaluated inherent aspects of the game such as clarity of objectives, feedback on player actions, and its closeness to educational concepts. The questions were separated into six dimensions:

- **Clarity of Objectives:** Evaluates whether the game is efficient in conveying its objectives.
- **Feedback:** Evaluates the quality and relevance of the information that the game provides to the player about their actions and progress.
- **Challenge Level:** Evaluates whether the game challenges are appropriate, progressive and motivating.
- **User Experience:** Evaluates whether the player finds the experience enjoyable and has reason to continue playing.
- **Perceived Control:** Evaluates whether the player feels that they have power and decision over their actions and progress.
- **Educational Concepts:** Evaluates how the game reinforces the concepts to be taught.

The *FORM 2: Game Evaluation Survey* contains 20 questions, with answers on a Likert scale from 1 to 5, where 1 means "Strongly Disagree" and 5 means "Strongly Agree", separated into the six dimensions described. It also contains an open response question where players could enter general comments. s

5 Results

Questionnaire responses were received in CSV format. They were processed using Microsoft Excel for data cleaning and subsequently analyzed using Google Colab with Python. Python tools were used to calculate descriptive statistics and generate representative graphs of the data. The questions were categorized into dimensions to facilitate data interpretation. Descriptive analyses were performed for each question and each dimension, including:

- Trend data such as averages and standard deviation to assess general trends in responses.
- Distribution of responses per question to visualize the frequency of each response.

No outliers were detected in the responses collected. Missing data were detected, but no action was required. There were two main cases of missing data in the collected data:

- Some students did not answer the optional qualitative question for additional comments. However, this question was not mandatory, so no steps were taken to handle these missing responses.

– Of the two surveys distributed, FORM 1 received 26 responses, while the FORM 2 received only 19. This suggests that some students completed only one of the surveys. Both surveys were analyzed independently, which minimized the impact of this discrepancy on the results.

A summary of the responses received in Form 1 can be seen in Table 2, and a summary of the responses received in Form 2 can be seen in Table 3.

Table 2. Trend Data, Form 1: REQ DUEL Usability Questionnaire

Dimension and Question	Average	Standard Deviation	Median	Mode
Dimension 1: UTILITY	**3.79**	**1.192**	**4**	**4**
1. The application helps me to be more effective.	4.04	0.958	4	4
2. The application helps me to be more productive.	3.85	0.967	4	4
3. I find the application useful.	4.31	0.970	5	5
4. The application gives me more control over my life activities.	3.54	1.272	4	4
5. The application makes the things I want to accomplish easier to do.	3.50	1.273	4	4
6. The application saves me time when I use it.	3.50	1.421	4	5
7. The application meets my needs.	3.65	1.263	4	4
8. The application does everything I would expect it to do.	3.96	1.216	4	5
Dimension 2: EASE OF USE	**4.17**	**1.069**	**4**	**5**
9. The application is easy to use.	4.38	0.941	5	5
10. The application is simple to use.	4.50	0.860	5	5
11. The application is user friendly.	4.46	0.905	5	5
12. The application requires as few steps as possible to achieve what I want.	4.38	0.941	5	5
13. The application is flexible.	4.00	0.938	4	4
14. The use of the application does not require much effort.	4.23	1.032	4.5	5
15. I can use the application without written instructions.	3.50	1.334	3.5	5
16. No inconsistencies are noted while using the application.	4.08	1.093	4	5
17. I think the application would appeal to both occasional and regular users.	3.81	1.297	4	5
18. I can always use the application successfully.	4.31	0.970	4.5	5
Dimension 3: EASE OF LEARNING	**4.38**	**0.939**	**5**	**5**
19. I learned to use the application quickly.	4.46	0.905	5	5
20. I easily remember how to use the application.	4.58	0.902	5	5
21. It is easy to learn how to use the application.	4.46	0.989	5	5
22. I have quickly become proficient with the application.	4.50	0.906	5	5
23. It is not necessary to explore new features by trial and error.	4.15	0.834	4	4
24. The order of the work areas is logical.	4.12	1.071	4	5
Dimension 4: SATISFACTION	**4.11**	**1.117**	**4**	**5**
25. I am satisfied with the application.	4.46	0.989	5	5
26. I would recommend the application to a friend.	3.92	1.230	4	5
27. The application is fun to use.	4.42	0.987	5	5
28. The application works the way I want it to work.	4.19	1.059	4	4
29. The application is wonderful.	3.81	0.939	4	4
30. I feel the need to have the application.	3.54	1.392	4	5
31. The application is pleasant to use.	4.42	0.902	5	5
TOTAL	**4.10**	**1.106**	**4**	**5**

It can be seen in these tables that, in general, all the results have an average close to 4, with a standard deviation close to 1. This indicates that the reception of the game was positive, with a low enough dispersion to consider the data as consistent.

Table 3. Trend Data, Form 2: Gaming Evaluation Survey.

Dimension and Question	Average	Standard Deviation	Median	Mode
Dimension 1: CLARITY OF OBJECTIVES	**4.58**	**0.680**	**5**	**5**
1. The objectives of the game were clear.	4.47	0.841	5	5
2. The objectives of each level of the game were clear.	4.68	0.582	5	5
3. The objectives of the game's puzzles were clear.	4.58	0.607	5	5
Dimension 2: FEEDBACK	**4.22**	**1.138**	**5**	**5**
4. Feedback on my progress in the game was adequate.	4.05	1.224	5	5
5. Feedback on my actions to solve the puzzles was adequate.	4.16	1.068	5	5
6. Feedback on the effect of my actions on the game was well-timed.	4.42	1.121	5	5
7. The feedback I received let me know my progress towards achieving the game's objectives.	4.26	1.195	5	5
Dimension 3: CHALLENGE LEVEL	**3.89**	**1.401**	**5**	**5**
8. The challenge level of the game suited my ability.	4.32	1.157	5	5
9. The challenge level varied as i progressed through the game.	3.79	1.653	5	5
10. The game provides new challenges with adequate progression.	3.74	1.522	4	5
11. The game provides puzzles with different levels of difficulty.	3.74	1.240	4	5
Dimension 4: USER EXPERIENCE	**4.49**	**0.947**	**5**	**5**
12. I have enjoyed the experience of playing the game.	4.68	0.749	5	5
13. It was satisfying to find out how to solve each of the game's challenges.	4.74	0.452	5	5
14. One of my motivations to continue playing was to solve the new challenges that came my way.	4.05	1.311	5	5
Dimension 5: PERCEIVED CONTROL	**4.09**	**1.184**	**5**	**5**
15. I have control over my actions and their impact on the game.	4.37	0.831	5	5
16. I feel that I have control over my success and progress in the game.	3.84	1.463	4	5
17. I feel that I can decide how to advance in the game.	4.05	1.177	4	5
Dimension 6: EDUCATIONAL CONCEPTS	**4.56**	**0.655**	**5**	**5**
18. The game mechanics relate to the concepts to be taught.	4.68	0.582	5	5
19. Game mechanics help exemplify the concepts that are intended to teach.	4.53	0.697	5	5
20. Participating in the game helps to understand the proposed concepts in a practical way.	4.47	0.697	5	5
TOTAL	**4.28**	**1.088**	**5**	**5**

Figures 2 and 3 show proportionally the number of positive (4 & 5), neutral (3) and negative (1 & 2) responses for each dimension in Forms 1 and 2 respectively. All dimensions have a percentage of at least 65% positive responses in both Forms.

Figure 2 shows that the dimension with the lowest percentage of positive responses in Form 1 was **utility**, with 69.2%. According to Table 2, several of the questions in this dimension obtained an average close to 3.5, the lowest average of all the questions. It can be concluded that the Utility value of the application is its weakest point.

On the other hand, the **ease of learning** dimension was the best evaluated, with a percentage of positive responses of 87.8%. It can be concluded that the application is sufficiently user-friendly so as not to make its use difficult. This idea is reinforced by the values of Median and Mode of this dimension, seen on Table 2, both with a value of 5, and its average of 4.38.

As for the second form, as can be seen in Fig. 3, the worst evaluated dimension was **challenge level**, with a percentage of positive responses of 67.1% and negative responses of 17.1%. Although this is not a low percentage, in relation to the rest of the percentages, it shows a weak point of the game. In addition, this dimension has an average response rate of 3.89, the lowest of all the dimensions of this form, as can be seen in Table 3.

The best evaluated dimensions, with percentages of positive responses very close to each other, are **clarity of objectives** with 93% of positive responses,

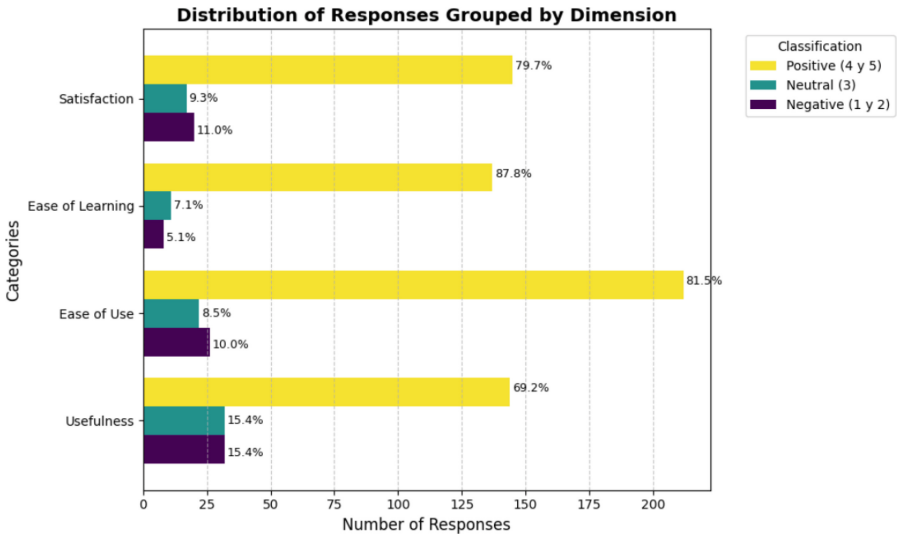


Fig. 2. Proportion of positive and negative responses per Dimension of Form 1

and **educational concepts** with 91.2%. The good evaluation in the clarity of objectives can be correlated with the high percentage that the **ease of learning** dimension obtained in Form 1: if the objectives are clear, one can learn to play easily as there are no obstacles to understanding what to do. As for the dimension of **educational concepts**, the high percentage of positive responses attests to the effectiveness of the game in teaching the concepts it intends to teach, which is one of its main objectives.

In addition to the quantitative data received as response scores, participants had the opportunity to provide feedback on their experience with the game.

The analysis of the open-ended comments provided by the participants allowed us to identify both strengths and areas for improvement related to the game. These comments complement the quantitative data and provide valuable information on user perception.

- Identified Strengths:
  - Easy to learn: Users mentioned that the game is intuitive and allows for quick immersion.
  - Entertaining: Participants rated the experience as entertaining and attractive.
  - Visual appeal: Some comments highlighted the design of the charts and the user-friendly interface.
- Areas for Improvement:
  - Dependence on luck: Several participants mentioned that success in the game is overly dependent on luck in drawing cards.
  - Limited variety: It is suggested to increase the diversity of cards and scenarios to maintain interest and offer more meaningful challenges.

- Additional functionalities: Improvements such as the option to cancel actions, restart levels and see the score in real time were proposed.

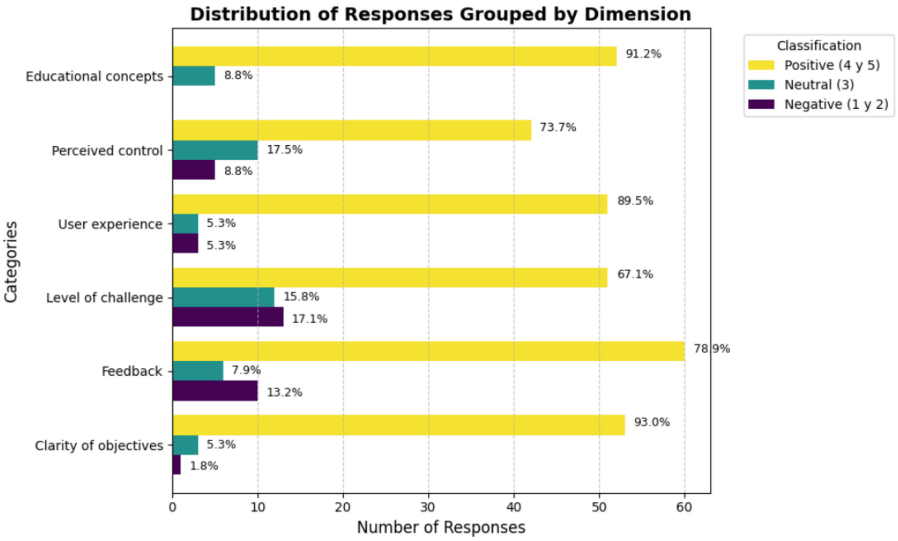


Fig. 3. Proportion of positive and negative responses per Dimension of Form 2

6 Discussion

The main hypothesis of the study was that the developed game would facilitate the learning and practice of fundamental concepts of Requirements Engineering, specifically in the areas of elicitation and analysis. The results obtained in the quantitative surveys partially support this hypothesis: it was observed that the dimensions related to user experience and clarity of objectives were positively evaluated by the participants, while the usefulness of the game as an educational tool obtained slightly lower ratings.

The qualitative analysis highlighted aspects such as the ease of use and the entertaining nature of the game, although areas for improvement were also identified, such as the dependence on luck and the need for more complex scenarios.

Several similarities can be found between the bibliography analyzed at the beginning of this document and the developed game proposal. ***Req Duel*** allows the player to take the role of someone who must exercise his or her skills to solve a problem within a simulation of a real situation. This aspect is like games such as *TREG*, *Biyubi*, *Requengin* and *BARA*, in which the player took the role of a Requirements Engineer to make bids.

On the other hand, like games such as *RE-O-Poly*, *ReGo* and *The Requirements Island*, ***Req Duel*** focuses on evaluating previously acquired knowledge,

but in a less direct way than those just mentioned. Most of the games studied obtained generally positive results in terms of their capacity as an educational tool. Likewise, ***Req Duel*** obtained positive results in the dimension of educational concepts, which was the main objective of this work.

As aspects to improve the game, there are three important points to consider: the **utility** and **challenge level** dimensions, and the frequent comments about the dependence on luck.

The **utility** dimension in Form 1 scored lower compared to other categories; this can be understood as a low capacity of the game to offer concepts extrapolated to students' lives. Questions such as *2. The app helps me be more productive* with an average of 3.85, and *5. The application makes the things I want to accomplish easier to do* with an average of 3.5 demonstrate that the skills acquired in the game were perceived as not very useful in a context other than this one. To improve this aspect of the game it would be necessary to reconsider the content of the cards to better illustrate the skills they are intended to reinforce. It may also be necessary to enrich the main mechanics of the game to better promote the critical thinking of the player, since in the current state, the game mechanics are very simple.

On the other hand, the relatively negative rating of the **challenge level** dimension of Form 2 can be understood by a lack of variety in its challenges. The game has little variability of levels in the game, as it has only one level always with the same collection of cards available. Questions such as *10. The game provides new challenges with adequate progression* and *11. The game provides puzzles with challenges of different difficulty levels*, both with an average score of 3.74, account for this weakness in the application. One solution would be to incorporate more cards so that the player can build their own strategies, and to add different levels with different enemy cards.

Lastly, the dependence on luck, repeatedly mentioned in the comments to the qualitative questions, might have influenced how some participants perceived the usefulness of the game, which may also explain the negative evaluation of this dimension. With a strong dependence on luck, the player perceives less control over the consequences of the game and less consistency in the results. With these obstacles, it is difficult to find any usefulness in the app.

In terms of notable aspects of the application, quantitative data shows a positive trend in key dimensions such as **ease of learning**, **clarity of objectives** and **educational concepts**, suggesting that the game succeeds in engaging students in a practical and entertaining environment.

Based on the good reception in the dimensions of **ease of learning** and **clarity of objectives**, it can be concluded that the game does a good job of explaining its mechanics and the conditions of victory and defeat. These aspects are taught to the player before starting a game and can be reinforced during the game with the help button. Although it is also possible to argue that the player profile shared by the students surveyed may have influenced the easy understanding of the game, since a large part of this group may already be familiar with card games from before.

The good reception in the **educational concepts** dimension shows that, despite the weaknesses in other dimensions, the game does a good job of reinforcing Requirements Engineering concepts that players have previously acquired in their educational process. Question *20. Participating in the game helps to understand the proposed concepts in a practical way* in Form 2 obtained an average of 4.47. This question illustrates the game's ability to reinforce this knowledge.

7 Conclusions

In relation to question **RQ1**, we can conclude that the game is motivating enough to keep students interested. The results obtained, both in quantitative surveys and in qualitative comments, indicate that it managed to capture and hold the attention of participants effectively. Dimensions such as Satisfaction and User Experience obtained a high percentage of positivity, showing that students found the game entertaining and dynamic.

Furthermore, many of the responses to the qualitative questions on the forms supported this positive perception. Also, during the playtesting, direct feedback of approval was received from the participants.

Regarding question **RQ2**, broadly speaking, the game performed well in most of the dimensions evaluated, obtaining a positivity percentage above 70% in all dimensions. The categories of Clarity of Objectives and Ease of Learning obtained the highest positivity, which shows that the participants easily understood the goals of the game and the mechanics necessary to achieve them. This demonstrates that the design fulfills its main purpose of offering a clear and accessible educational environment.

However, the dimensions of Usefulness and Level of Challenge scored lower in comparison, although still positive. This suggests that while the game is functional, it could benefit from more complex and realistic scenarios that present significant challenges. These adjustments to the complexity of game mechanics would allow students to fully apply their requirements engineering skills. The alignment between gameplay and domain knowledge, could be further tested comparing student gameplay logs with scores in a requirement engineering test, to prove the usefulness of the game as a teaching tool.

7.1 Limitations

The study presented certain challenges that may influence the interpretation of the results. Although the two surveys were administered to the same group of students, there was a discrepancy in the number of responses obtained (26 in Form 1 and 19 in Form 2). This difference could be attributed to the level of interest or willingness of the participants to respond to both forms. However, since the surveys were analyzed independently, this discrepancy should not significantly affect the conclusions drawn from each instrument.

Regarding the sample size, with 26 participants, it is important to highlight that this is within the expected ranges for an exploratory analysis, such as the

one carried out in this study. Although a larger group could provide greater robustness to the findings, the data obtained are valid for generating initial hypotheses about the effectiveness of the game as an educational tool.

The results obtained in this study are specific to students who are learning about Requirements Engineering. Since the game is designed as an educational tool for students who have not yet put these skills into practice, it is not considered to be useful for professionals who are already working in this field or for common users without formal knowledge of Software Engineering. Although the game turned out to be entertaining, its educational value is focused on providing a practical experience to evaluate skills that students are in the process of developing.

7.2 Future Works

It is recommended to carry out future studies with larger and more diverse groups to confirm these findings and explore their applicability in different contexts. This would allow us to evaluate whether the conclusions obtained in this study remain consistent as the scope of the sample is expanded, which would contribute to validating the findings in a more robust manner.

Based on the feedback received, several areas for improvement were identified for the game. Among the most relevant ones, improvements to the game's quality of life stand out, such as the implementation of options to cancel or restart levels, as well as a more interactive interface. In addition, the game could benefit from additional levels with a higher level of difficulty, which would provide a more challenging experience for students.

However, the most significant improvement lies in making the scenarios more complex and realistic. Currently, the game features a dynamic where students must solve stakeholder problems using solution cards. If this dynamic were deepened with more complex scenarios and situations that represent real challenges, the game could offer a more effective educational challenge, promoting a true decision-making exercise and practice of Requirements Engineering skills.

References

1. Abt, C.: Serious Games. University Press of America (1987)
2. Alexander, M., Beatty, J.: Effective design and use of requirements engineering training games. In: 2008 Requirements Engineering Education and Training, pp. 18–21. IEEE (2008)
3. Bellotti, F., Kapralos, B., Lee, K., Moreno-Ger, P., Berta, R.: Assessment in and of serious games: an overview. Adv. Hum. Comput. Interact. **2013**(1), 136864 (2013)
4. Brandl, L.C., Schrader, A.: Serious games in higher education in the transforming process to education 4.0-systematized review. Educ. Sci. **14**(3), 281 (2024)
5. Daun, M., Grubb, A.M., Stenkova, V., Tenbergen, B.: A systematic literature review of requirements engineering education. Requirements Eng. **28**(2), 145–175 (2023)
6. Dick, J., Hull, E., Jackson, K.: Requirements Engineering. Springer (2017)

7. Garcia, I., Pacheco, C., León, A., Calvo-Manzano, J.A.: Experiences of using a game for improving learning in software requirements elicitation. Comput. Appl. Eng. Educ. **27**(1), 249–265 (2019)
8. Garcia, I., Pacheco, C., León, A., Calvo-Manzano, J.A.: A serious game for teaching the fundamentals of ISO/IEC/IEEE 29148 systems and software engineering-lifecycle processes-requirements engineering at undergraduate level. Comput. Stand. Interfaces **67**, 103377 (2020)
9. Girard, C., Ecalle, J., Magnan, A.: Serious games as new educational tools: how effective are they? a meta-analysis of recent studies. J. Comput. Assist. Learn. **29**(3), 207–219 (2013)
10. Hainey, T., Connolly, T.M., Stansfield, M., Boyle, E.A.: Evaluation of a game to teach requirements collection and analysis in software engineering at tertiary education level. Comput. Educ. **56**(1), 21–35 (2011)
11. Ibrahim, Z., Soo, M.C., Soo, M.T., Aris, H.: Design and development of a serious game for the teaching of requirements elicitation and analysis. In: 2019 IEEE International Conference on Engineering, Technology and Education (TALE), pp. 1–8. IEEE (2019)
12. Jaramillo, C.: Communication and traceability game: a way to improve requirements elicitation process teaching. Revista Facultad de Ingeniería Universidad de Antioquia **56**, 213–221 (2010)
13. Knauss, E., Schneider, K., Stapel, K.: A game for taking requirements engineering more seriously. In: 2008 Third International Workshop on Multimedia and Enjoyable Requirements Engineering-Beyond Mere Descriptions and with More Fun and Games, pp. 22–26. IEEE (2008)
14. Lamb, R.L., Annetta, L., Firestone, J., Etopio, E.: A meta-analysis with examination of moderators of student cognition, affect, and learning outcomes while using serious educational games, serious games, and simulations. Comput. Hum. Behav. **80**, 158–167 (2018)
15. Liang, P., De Graaf, O.: Experiences of using role playing andwiki in requirements engineering course projects. In: 2010 5th International Workshop on Requirements Engineering Education and Training, pp. 1–6. IEEE (2010)
16. Lima, T., Campos, B., Santos, R., Werner, C.: Ubire: a game for teaching requirements in the context of ubiquitous systems. In: 2012 XXXVIII Conferencia Latinoamericana En Informatica (CLEI), pp. 1–10. IEEE (2012)
17. Liu, Y., Li, T., Huang, Z., Yang, Z.: Bara: a dynamic state-based serious game for teaching requirements elicitation. In: 2023 IEEE/ACM 45th International Conference on Software Engineering: Software Engineering Education and Training (ICSE-SEET), pp. 141–152. IEEE (2023)
18. Quintanilla Portugal, R.L., Engiel, P., Pivatelli, J., do Prado Leite, J.C.S.: Facing the challenges of teaching requirements engineering. In: Proceedings of the 38th International Conference on Software Engineering Companion, pp. 461–470 (2016)
19. Rice, J.: Assessing higher order thinking in video games. J. Technol. Teach. Educ. **15**(1), 87–100 (2007)
20. Rosas, R., et al.: Beyond nintendo: design and assessment of educational video games for first and second grade students. Comput. Educ. **40**(1), 71–94 (2003)
21. dos Santos, S.C., Soares, F.S.: Authentic assessment in software engineering education based on PBL principles a case study in the telecom market. In: 2013 35th International Conference on Software Engineering (ICSE), pp. 1055–1062. IEEE (2013)

22. Shute, V.J., Ventura, M., Bauer, M., et al.: Melding the power of serious games and embedded assessment to monitor and foster learning: flow and grow. In: Serious games, pp. 317–343. Routledge (2009)
23. Smith, R., Gotel, O.: Gameplay to introduce and reinforce requirements engineering practices. In: 2008 16th IEEE International Requirements Engineering Conference, pp. 95–104. IEEE (2008)
24. Thiry, R., Marcello, G.: Development of a game to support the teaching of requirements engineering: the requirements island. Proc. SBGames SBC Florianópolis, Brazil 358–361 (2010)
25. Vega, K., Fuks, H., Carvalho, G.: Training in requirements by collaboration: branching stories in second life. In: 2009 Simposio Brasileiro de Sistemas Colaborativos, pp. 116–122. IEEE (2009)
26. Wouters, P., Van Nimwegen, C., Van Oostendorp, H., Van Der Spek, E.D.: A meta-analysis of the cognitive and motivational effects of serious games. J. Educ. Psychol. **105**(2), 249 (2013)
27. Zhonggen, Y.: A meta-analysis of use of serious games in education over a decade. Int. J. Comput. Games Technol. **2019**(1), 4797032 (2019)

Identification of Older Adults' Characteristics that Affect the Usability of Mobile Applications: A Tertiary Study

John W. Castro[1(✉)], Agustín I. Astudillo[2], and Luis Rojas[3]

[1] Departamento de Ingeniería Informática y Ciencias de la Computación, Universidad de Atacama, Copiapó, Chile, y Centro Interuniversitario de Envejecimiento Saludable (CIES), Santiago, Chile
john.castro@uda.cl

[2] Departamento de Ingeniería Informática y Ciencias de la Computación, Universidad de Atacama, Copiapó, Chile
agustin.astudillo.17@alumnos.uda.c

[3] Facultad de Ingeniería, Universidad San Sebastián, Bellavista 7, 8420524 Santiago, Chile
luis.rojasp@uss.cl

Abstract. Considering the growth of the older adult population and technological advancements, this age group faces challenges when interacting with mobile applications, as these applications often fail to consider age-specific characteristics during their design. Some primary studies identify older adults' characteristics that may affect the usability of mobile applications, as well as secondary studies. However, to the best of our knowledge, no tertiary study has conducted a global analysis to identify age-related characteristics that may impact the usability of mobile applications. For this reason, our tertiary study aims to determine the characteristics of older adults that can influence the usability of mobile applications. We searched for secondary studies published between 2018 and 2024 using the Scopus, Web of Science, and IEEE Xplore databases. We identified eight secondary studies focused on these characteristics and their impact on usability, grouping them into seven categories: (i) psychomotor, (ii) cognitive, (iii) visual, (iv) auditory, (v) motivational, (vi) psychosocial, and (vii) economic. We conclude that it is necessary to adapt usability evaluation techniques to suit older adults' needs better.

Keywords: Usability · older adult · smartphone · ICT · aging barriers · tertiary study · HCI

1 Introduction

In recent years, there has been a significant increase in the global older adult population [1]. By 2025, 22% of the world's population will be over 60; by 2030, one in six people worldwide will be 60 years old or older [1]. Meanwhile, in recent decades, new technologies have advanced and accelerated [2]. Information and communication

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 37–56, 2025.
https://doi.org/10.1007/978-3-031-93539-8_3

technologies (ICTs), such as online applications, have become essential to our lives [3]. For older adults, ICTs can be a tool to improve their quality of life [4]. Some benefits this age group can gain from using ICTs include the ease of communicating with family members, which helps reduce loneliness [5], or strengthening their social relationships and intergenerational bonds [6].

The smartphone is the most commonly used technological device among older adults, mainly for communication, social networking, entertainment, and health-related applications [7]. However, this age group is often excluded due to technology or software systems not being designed with their needs in mind. Typically, during the development of these systems, age-related characteristics such as cognitive decline and visual impairments are not considered [8-10]. These age-specific characteristics pose challenges in the use of mobile applications [11] and may affect their usability [12].

Usability is the extent to which specific users can use a system, product, or service to effectively, efficiently, and satisfactorily achieve specific objectives in a particular context of use [13]. Moreover, usability comprises attributes such as ease of learning, efficiency, error rate, and user retention over time [14]. On the other hand, user experience refers to the perceptions and responses of a person resulting from the use and/or anticipated use of a product, system, or service [15]. Since our study focuses on understanding older adults' characteristics that may affect the usability of mobile applications, we prioritize usability over user experience.

Some challenges older adults face due to their characteristics include difficulty recognizing visual elements or interacting with moving elements in a graphical user interface [12]. Identifying these characteristics is essential for those developing mobile applications, ensuring designs are adapted to these needs, thereby improving usability.

Its usability must be evaluated to determine how usable an application or software system is. Usability evaluation is conducted through techniques such as the "think aloud" protocol [14], which allows researchers to study participants' mental processes as they verbalize their thoughts while performing a task [16]. Some authors argue that usability evaluation techniques must be adapted to align with older adults' characteristics [17-19], as applying them in their original form presents challenges when involving this group. For example, factors such as cognitive decline can influence the application of the "think aloud" technique. Therefore, understanding these characteristics is vital to adaptations to usability evaluation techniques.

Some primary studies address older adults' characteristics [20-25]. For instance, in the study by Gao et al. [20], the authors report that this age group has age-related psychomotor deficiencies, such as longer response times in complex motor tasks, which affect their ability to perform interactions requiring repetitive and rapid movements. Another primary study by Kascak et al. [21] reports that with aging, capabilities such as vision, motor skills, cognitive abilities, and hearing decline. On the other hand, secondary studies have compiled these findings. For example, in the secondary study by Elguera et al. [26], the authors identify the main usability challenges older adults face when interacting with a graphical user interface and explore how such applications can be more acceptable for this population. In the study by Kurniawan et al. [27], the authors report on another secondary study to identify the challenges older adults face when

using mHealth applications and the usability dimensions most commonly employed to evaluate mHealth apps.

Given the above, conducting a more extensive characterization in this context is more crucial than a systematic literature review can provide. Moreover, based on preliminary research, no study offers a comprehensive overview of existing literature reviews related to age-associated characteristics of older adults that may affect the usability of mobile applications. Therefore, it is necessary to develop a specific tertiary study that provides researchers and developers with a unified perspective on this topic, serving as a high-level catalog of existing research and consolidating findings in this area.

This study is organized as follows: Sect. 2 presents related work. Section 3 details the research methodology. Section 4 reports and discusses the results. Section 5 addresses threats to validity, and finally, Sect. 6 presents the conclusions and future work.

2 Related Work

In the initial stage of our research, we identified a single study [28] that reviewed primary and secondary studies. In the work by Wildenbos et al. [28], the authors report the findings of a scoping review focused on understanding the barriers that affect the usability of mobile health applications for older adults and propose a usability-centered framework for mHealth tailored to older users. The systematic literature review utilized the PUBMED, EMBASE, ScienceDirect, and Web of Science databases. The authors identified four categories of barriers: (i) cognitive, (ii) physical, (iii) perceptual, and (iv) motivational. They also noted that these barriers could limit the utility and usability of mHealth for older adults, creating difficulties such as holding a mobile device with one hand or interacting with buttons placed too closely together.

Additionally, we identified the literature review conducted by Elguera et al. [26]. The authors performed a systematic literature review covering 2013 to 2019 using the IEEE Xplore, Scopus, ProQuest Computing, and ACM Digital Library databases. Their study addressed the main usability challenges older adults face when interacting with graphical user interfaces in mobile applications. The primary findings highlighted three limitations: (i) visual, (ii) psychomotor, and (iii) cognitive.

In the work by Wildenbos et al. [28], the authors identified age-related barriers impacting user experience. However, their study focuses exclusively on barriers to mHealth applications. While the authors did not conduct a tertiary study, they included secondary studies, which is why we consider it related work. Moreover, the secondary study by Elguera et al. [26] identifies general problems related to older adults' limitations when interacting with mobile applications. However, it does not aim to identify the characteristics of older adults that might influence the usability of mobile applications.

For this reason, research, such as the present study, is necessary to contribute to this field by identifying the characteristics of older adults that could influence the usability of mobile applications. This will benefit developers of such applications by enabling designs that better accommodate these characteristics. It will also allow usability evaluators to adapt usability evaluation techniques to facilitate better participation of older adults. Ultimately, this will benefit older adults by providing applications better suited to their needs.

3 Research Methodology

This study is a tertiary investigation aimed at synthesizing and analyzing existing evidence on age-related characteristics of older adults that may affect the usability of mobile applications. The methodology follows a series of phases established by Kitchenham et al. [29]: (i) Formulating the research questions, (ii) defining the search strategy, (iii) establishing inclusion and exclusion criteria, (iv) selecting the studies, (v) extracting information relevant to the research questions, and finally (vi) synthesizing the extracted information.

3.1 Research Questions

This research focuses on three aspects: the characteristics of older adults, the use of mobile applications, and their usability. Consequently, the information extracted from secondary studies aims to answer the following research questions (RQ):

(RQ1): What are the characteristics of older adults that influence or impact the usability of mobile applications?

(RQ2): What future research directions are associated with usability from the perspective of older adult users?

3.2 Definition of the Search Strategy

The tertiary study begins by identifying the keywords that will be used to form the search string. To achieve this, we defined a Control Group (CG), which consists of an initial set of studies relevant to the research area of interest [30]. In this case, the focus is on the characteristics of older adults associated with aging that may affect the usability of mobile applications. The studies included in the CG must adequately address the proposed research questions [30].

The CG was identified to maintain objectivity when selecting the search string. To define the CG, we manually searched for studies relevant to the research context, ensuring they answered the proposed questions. As a result of this process, we identified five secondary studies for the CG [26, 28][31-33] and verified that these studies were indexed in Scopus, as it is the database with the highest number of studies. After verification, we confirmed that all CG studies were in Scopus. Following this, we extracted the keywords that would form the search string using Atlas.ti.

From the extracted keywords, we created a table listing the frequencies of all words and combinations related to the research objective found in each CG study. Initially, we selected only the terms related to the research questions and presented a significant percentage (at least 60%) of the CG studies. This coverage percentage was calculated using Eq. 1:

$$Coverage\ Percentage = \frac{\text{(Frequency of Appearance)x100}}{n} \quad (1)$$

where: (i) *Frequency of appearance* represents the number of studies within the CG that contain the word, and (ii) n is the total number of studies in the CG. Subsequently, we determined the weight of each word using Eq. 2.

$$Weight = \frac{\left(\frac{\text{coverage percentage}}{\text{maximum coverage}}\right) + \left(\frac{\text{word frequency}}{\text{maximum frequency}}\right)}{2} \tag{2}$$

where: (i) *Maximum coverage* is the highest possible coverage, corresponding to 100%, (ii) *word frequency* is the number of times the word appears across all studies in the CG, and (iii) *maximum frequency* is the highest frequency observed, which in this case was 501.

Table 1 presents the list of words obtained during this process, showing the total frequency of each word, the coverage percentage, and the respective weight. Regarding the values obtained for the weight variable, in the final stage, we only considered words with a weight equal to or greater than 0.41, ensuring a coverage percentage of at least 80% of the studies in the CG. Additionally, we generated seven search strings, as creating more combinations did not significantly reduce search results while maintaining the inclusion of the studies in the CG.

3.3 Inclusion and Exclusion Criteria

The selection criteria were defined as follows:

Inclusion Criteria. The study must:

1. Be written in English; AND
2. Consider older adults aged 60 years or older; AND
3. Report how the usability of mobile applications is influenced by the barriers, limitations, or characteristics specific to this age group, or report which characteristics of older adults affect the use of such applications.

Exclusion Criteria. The study:

4. Is not a secondary study; OR
5. Does not consider older adults as the study population; OR
6. Does not report how the usability of mobile applications is influenced by the barriers, limitations, or characteristics of this age group; OR
7. Does not report the characteristics of older adults that affect the use of mobile applications.

3.4 Selection Process

Once the keywords were identified, we proceeded to build the search strings. We identified their main components: (i) Characteristics, (ii) older adults, (iii) usability, (iv) mobile applications, and (v) systematic literature reviews. For the first four components, the terms to be used were obtained from Table 1. It is essential to mention that for the last component, corresponding to systematic literature reviews, we extracted the terms used from the study by Napoleao et al. [34], which were part of all the search strings

tested. Additionally, we added the term "Scoping Review" to this last component, as we identified in the literature that it is increasingly used to refer to studies conducting systematic literature reviews. We used the logical operator AND to connect each component and the logical operator OR to include synonyms. Considering the above, we created and tested seven strings with different combinations. In Table 2, we present the identifier (ID) for each string, the number of studies returned by the search in Scopus, the search string used, the number of studies found from the CG, X ratio, Y ratio, and the average.

Table 1. Word List Obtained from the CG Analysis.

Word	Frequency	Coverage percentage	Weight
elderly	501	100	1
usability	474	100	0.97
older	387	100	0.89
adults	254	100	0.75
user	215	100	0.71
mobile	202	100	0.70
ehealth	297	80	0.70
barriers	147	100	0.65
applications	125	100	0.62
app	91	100	0.59
apps	88	100	0.59
aging	78	100	0.58
application	76	100	0.58
computer	71	100	0.57
issues	62	100	0.56
devices	60	100	0.56
problems	50	100	0.55
human	39	100	0.54
interaction	36	100	0.54
senior	35	100	0.53
limitations	33	100	0.53
experience	30	100	0.53
smartphone	26	100	0.53
challenges	101	80	0.53
barrier	27	80	0.43

(*continued*)

Table 1. (*continued*)

Word	Frequency	Coverage percentage	Weight
adult	18	80	0.42
smartphones	15	80	0.41
experiences	13	80	0.41
characteristics	6	80	0.41
seniors	18	60	0.32
problem	8	60	0.31
hci	5	60	0.30

As shown in Table 2, all the strings found the five studies in the CG except one. To select the most appropriate string, we calculated the X and Y proportions and the average of these two. We used Eqs. 3, 4, and 5 to calculate each proportion.

$$X_{ratio} = \frac{\text{number of articles found in CG}}{\text{total articles in CG}} \tag{3}$$

$$Y_{ratio} = \frac{\text{number of articles found in CG}}{\text{total articles found by search string}} \tag{4}$$

$$Average = \frac{X_{ratio} + Y_{ratio}}{2} \tag{5}$$

The string with ID 5 was selected for the search because it returned the smallest number of studies and found all the studies in the CG. Therefore, it had the highest average proportion. After that, we searched for studies in Scopus, Web of Science (WoS), and IEEE Xplore. In Scopus, the search field we used was "Title-Abs-Key." We used the "Topic Search" field for WoS, while for IEEE Xplore, we used the "Abstract" field. In total, we found 369 studies. The search period was from 2018 to 2024. This period was chosen to analyze studies from the last six years and observe the evolution and trends in the field. After excluding duplicate studies, the number decreased to 297. We then applied the inclusion and exclusion criteria to the title and abstract of each non-duplicate study, reducing the number to 14 pre-selected studies. Finally, we applied the inclusion and exclusion criteria to each pre-selected study. This required downloading and reading each study thoroughly, reducing the number to 8 selected studies. Figure 1 shows the study selection process. The complete information on the secondary studies can be found in Appendix A.

Table 2. Tested Search Strings.

ID	Studies	String	CG	X_{ratio}	Y_{ratio}	Average
1	230	(barriers OR problems OR issues OR limitations OR challenges OR barrier OR characteristics) AND (elderly OR older OR "older adults" OR aging OR senior OR "older adult") AND (usability OR "human-computer interaction" OR "user experience" OR "user experiences") AND (mobile OR ehealth OR "mobile applications" OR "mobile aplication" OR "mobile devices" OR apps OR app OR smartphone OR smartphones) AND ("systematic review" OR "literature review" OR "systematic mapping" OR "mapping study" OR meta-analysis OR survey OR "literature analysis" OR "scoping review")	5	1	0,022	0,511
2	206	(barriers OR problems OR issues OR limitations OR challenges OR barrier) AND (elderly OR older OR "older adults" OR aging OR senior) AND (usability OR "human-computer interaction" OR "user experience") AND (mobile OR ehealth OR "mobile applications" OR "mobile devices" OR apps OR smartphones) AND ("systematic review" OR "literature review" OR "systematic mapping" OR "mapping study" OR meta-analysis OR survey OR "literature analysis" OR "scoping review")	5	1	0,024	0,512
3	122	(barriers OR problems OR issues OR limitations) AND (elderly OR older OR "older adults" OR aging) AND (usability OR "human-computer interaction") AND (mobile OR ehealth OR "mobile applications" OR "mobile devices" OR apps) AND ("systematic review" OR "literature review" OR "systematic mapping" OR "mapping study" OR meta- analysis OR survey OR "literature analysis" OR "scoping review")	4	0.8	0,033	0,416
4	174	(barriers OR problems OR issues OR limitations OR challenges OR barrier) AND (elderly OR older OR "older adults" OR aging OR senior) AND (usability OR "human computer interaction" OR "user experience") AND ("mobile applications" OR ehealth OR "mobile devices" OR apps OR smartphones) AND ("systematic review" OR "literature review" OR "systematic mapping" OR "mapping study" OR meta-analysis OR survey OR "literature analysis" OR "scoping review")	5	1	0,029	0,514

(*continued*)

Table 2. (*continued*)

ID	Studies	String	CG	X_{ratio}	Y_{ratio}	Average
5	153	(barriers OR problems OR issues OR limitations OR challenges OR barrier) AND (elderly OR older OR "older adults" OR aging OR senior) AND (usability) AND (mobile OR ehealth OR "mobile applications" OR "mobile devices" OR apps OR smartphones) AND ("systematic review" OR "literature review" OR "systematic mapping" OR "mapping study" OR meta-analysis OR survey OR "literature analysis" OR "scoping review")	5	1	0,033	0,516
6	206	(barriers OR problems OR issues OR limitations OR challenges) AND (elderly OR older OR "older adults" OR aging OR senior) AND (usability OR "human-computer interaction" OR "user experience") AND (mobile OR ehealth OR "mobile application" OR "mobile devices" OR apps OR app OR "mobile application" OR smartphone) AND ("systema tic review" OR "literature review" OR "systematic mapping" OR "mapping study" OR meta-analysis OR survey OR "literature analysis" OR "scoping review")	5	1	0,024	0,512
7	165	(barriers OR problems OR issues OR limitations OR challenges) AND (elderly OR older OR "older adults" OR aging OR senior) AND (usability OR "human-computer interaction") AND (mobile OR ehealth OR "mobile applications" OR app OR apps) AND ("systematic review" OR "literature review" OR "systematic mapping" OR "mapping study" OR meta-analysis OR survey OR "literature analysis" OR "scoping review")	5	1	0,030	0,515

4 Results and Discussion

In this section, we report the results obtained from our tertiary study, synthesizing the secondary studies analyzed and answering the research questions. In Fig. 2, we summarize the information obtained from the identified secondary studies. This figure is divided into two areas that provide a global view of the results. On the left are two XY scatter plots with bubbles, representing the intersections of the type-year of publication set and the publication type-characteristics categories. The size of the bubbles represents the number of secondary studies classified in each set. On the upper right side, we observe that 2021 had the highest number of publications, with a decrease in the number of publications in 2022 and 2023, followed by a rebound in 2024.

From now on, the research questions posed in this tertiary study will be answered.

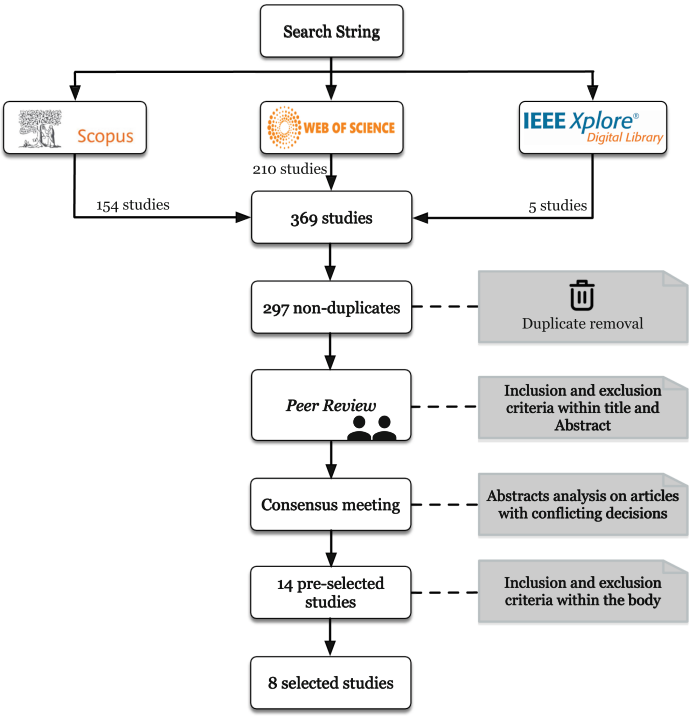


Fig. 1. Study selection process.

4.1 Characteristics of Older Adults

In this section, we answer the research question RQ1: What characteristics of older adults can influence or impact the usability of mobile applications? Based on the analysis of the eight secondary studies, we identified seven categories of characteristics, which correspond to limitations: (i) Psychomotor, (ii) Cognitive, (iii) Visual, (iv) Auditory, (v) Motivational, (vi) Psychosocial, and (vii) Economic.

Psychomotor Limitation. For this category, we identified and synthesized the following characteristics:

- Increased time needed to complete tasks while using a smartphone. There is a general rule that older adults require between 50% to 100% more time to complete a task while using smartphones compared to individuals under 30 years old [ES1][ES6]. This time may be due to the need to read the information on a small screen, as in the case of smartphones [ES4].
- Reduced motor skills. This presents a challenge for older adults when using smartphones and can make it more difficult to interact with small buttons on an interface [ES6].

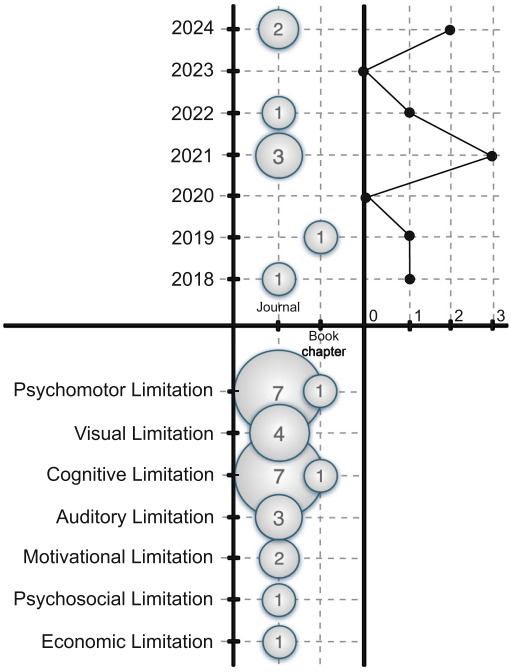


Fig. 2. Mapping of distribution of secondary education between categories of older adult characteristics and the type of publication.

- Slower movements and reflexes. This condition can affect learning time, execution speed, and failure rate [ES2]. Additionally, older adults prefer that all icons be on a single screen to avoid scrolling [ES2].
- Rheumatoid arthritis. This disease can cause discomfort when holding a device [ES2][ES8], limiting its utility and usability [ES8].
- Tighter muscles and joints and decreased balance. Older adults with this characteristic have difficulty navigating an interface and are affected by slower execution speeds [ES2].
- Difficulty entering data with fingers. Refers to the problems older adults face when entering data using their fingers [ES3].
- Trembling hands. Older adults experience hand tremors during the use of technology [ES5]. Additionally, their movements and reflexes are slower [ES8].
- Loss of muscle mass and increased variability of fine motor movements. Reduced psychomotor activity can be measured by the loss of muscle mass and increased variability of fine motor movements [ES6].
- Reduced grip strength. This can cause difficulty navigating menus if the smartphone is held with one hand [ES6].
- Decreased hand-eye coordination. This characteristic causes problems when performing gestures to move objects or zoom, as older adults cannot coordinate their fingers well [ES6].

- Less precise control of movement. Causes issues when interacting with poorly positioned buttons [ES6].
- Decreased hand movement retention can cause problems for older adults when interacting with a scrollbar at the top or right side [ES6].

Cognitive Limitation. For this second category, we identified and synthesized the following characteristics:

- Decreased memory capacity. Memory capacity significantly decreases with age, impacting the ability to recall planned actions [ES1]. Older adults also tend to forget actions they intend to carry out [ES6].
- Difficulty remembering actions and information. While using applications, older adults may repeatedly forget steps or actions required at the moment [ES3], such as available functions within an app [ES5]. They also have difficulty remembering tasks that must be completed in the future [ES6]. Additionally, older adults struggle remembering passwords or new information [ES2].
- Decreased work capacity and concentration. This is due to the reduced ability of older individuals to process information over time. It is also linked to a decrease in concentration, which negatively affects the use of mobile applications [ES2][ES8].
- Requiring more time to complete a task. This characteristic overlaps with psychomotor limitation, which some authors categorize as a cognitive barrier. Older adults need more time to complete a task and find it much more challenging to do so effectively [ES2].
- Difficulty with large amounts of information. Older adults can feel overwhelmed or stressed by excessive amounts of information [ES2].
- Lack of experience and understanding of terms. Older adults lack expertise and a sense of terminology or jargon used by younger generations. Therefore, it is suggested that a familiar language with direct meaning be created [ES2].
- Confusion. Related to the lack of knowledge of terms, older adults can become confused by unfamiliar terms, as processing these unknown terms requires the brain to process more information [ES4].
- Difficulty learning new skills. Older adults need to dedicate more time and effort to learning new skills [ES2][ES6][ES8].
- Decreased reasoning, semantic fluency, and spatial cognition. Reduced reasoning can cause complexity in understanding graphics or difficulty recognizing icons. Reduced semantic fluency can create problems in understanding terms. Spatial cognition can lead to disorientation when interacting with moving menus and transition effects [ES6].
- Decreased cognitive function. A decrease in cognitive function increases the effort needed to learn new skills, such as operating new technologies [ES7].
- Decreased processing speed refers to the reduced speed at which information is processed, which forces older adults to rely on a theory-based approach for decision-making [ES7].

Visual Limitation. In this third category, we identified and synthesized the following characteristics:

- Decreased visual capacity. Older adults have worse vision compared to younger individuals [ES1]. Around 15.8% of adults between 65 and 74 years old have vision problems affecting their daily lives [ES2].
- Light-related issues. Older adults' eyes are slower to accommodate dark and light spaces. They also react slower to quick changes in light intensity [ES1]. Additionally, to see more clearly, they need additional lighting [ES2][ES8], so they must configure a screen with appropriate brightness [ES2].
- Decreased focus ability. Older adults cannot quickly change focus [ES1]. Furthermore, the ability to focus on objects up close or far away decreases with age [ES2].
- Reduced contrast sensitivity. Older adults experience contrast sensitivity, affecting the combination of colors and the readability of text over colored backgrounds [ES2]. They also have difficulty distinguishing between contrasts, especially in low-light settings [ES6].
- Macular degeneration and bifocal lenses. Macular degeneration is considered a vision condition [ES2][ES4]. This characteristic refers to vision loss caused by damage to the retina, where a shadow appears in the center of the visual field [ES2]. Vision impairment is often associated with bifocal lenses [ES2].
- Eye fatigue. Due to eye fatigue, older adults find it challenging to use smartphone screens [ES2].
- Poor readability visibility. Older adults have trouble reading information on small mobile phone screens. This relates to the increased time needed to complete a task, influenced by reading difficulties [ES4].
- Decreased visual sharpness. Older adults have lower visual acuity, which describes the clarity or sharpness of vision and can be determined under different lighting conditions. This can affect the identification of visual elements such as images, graphics, text, icons, and buttons [ES6].
- Decreased color vision. Older adults struggle to distinguish colors, particularly violet, blue, and green [ES6].
- Decreased adaptation to darkness. With aging, older adults have less ability to adapt to darkness. They need more light to see clearly [ES6].
- Decreased visual function. Refers to the reduced ability to see objects in detail, such as distinguishing colors [ES6].

Auditory Limitation. For this fourth category, we identified and synthesized two characteristics. First, older adults face challenges distinguishing short-duration acoustic signals and find it difficult to understand imprecise feedback [ES6]. In general, older adults with moderate to severe hearing difficulties use the Internet less [ES8] or perform worse when using an application than those without these conditions [ES6]. Second, older adults are more likely to suffer from deafness or have hearing impairment [ES2][ES6].

Motivational Limitation. For this fifth category, we identified and synthesized two characteristics. The first is the desire to stop using an application if the benefits are unclear or insignificant [ES6][ES8]. The second is the lack of confidence older adults often have in using mobile phones. They fear experimenting with these devices, fearing they might break something [ES8].

Psychosocial Limitation. In this sixth category, we identified and synthesized three characteristics. First, older adults may encounter applications incompatible with their lifestyle goals [ES6]. Second, older adults with experience using similar devices may have doubts or lack knowledge of possible uses when using mobile applications [ES6]. Third, when adopting new technologies, older adults are often more conservative, skeptical, cautious, traditional, and distrustful of innovation. They are also less willing to use technology they consider less valuable or more complicated [ES6].

Economic Limitation. In the final category, we identified that older adults are more concerned about the costs of devices and internet connections. Their pensions may be limited, resulting in lower incomes than when they were actively working. This can prevent them from using specific devices if their costs exceed what they can afford [ES4].

Table 3 summarizes each characteristic identified by category and indicates the secondary study where it was recognized. The characteristics of each category will be detailed next.

As we saw earlier, various characteristics of older adults can influence the usability of mobile applications. However, since specific usability evaluation techniques require, for example, that the user of an application perform a task provided by the evaluator, and considering that older adults need between 50% and 100% more time to complete tasks with smartphones compared to younger individuals, the time allocated to perform these tasks must be extended. Thus, usability evaluation techniques involving older adults' participation require such adaptations.

4.2 Future Research of Secondary Studies

In this section, we address the research question RQ2: What future research directions are associated with usability from the perspective of older adult users? We identified five main future works in the secondary studies. The first is the recommendation for developers. The authors of [ES1] provide recommendations based on Android and iOS app development guidelines for issues caused by visual, psychomotor, and cognitive limitations. The goal is to create better applications that fit the conditions of this age group. Second, there is a need to research technology that helps increase the adoption of mHealth apps, as an adapted app will make it more accessible for this age group [ES2]. Additionally, new interactive methods must be explored to make mHealth apps more usable and attractive to older adults. It is also recommended that the suggestions of older adult users be considered when designing applications that are easy to use for this group.

Third, usability evaluation techniques need to be modified. In the study [ES3], the authors point out that usability evaluation techniques like joint evaluation, cooperative usability testing, and think-aloud cannot be used in their original versions. Therefore, there is a need to explore and propose evolved versions of these usability evaluation methods [ES3]. It is also relevant to explore how older adults behave during usability evaluations [ES3]. Fourth, it is essential to examine the barriers and opportunities for older adults to access and use technology [ES5]. Finally, efforts are needed to investigate how cognitive, physical, and perceptual barriers are related to older adults' disinterest in using mHealth apps [ES8].

Table 3. Identified characteristics by category.

Category	Identified Characteristic	Secondary Study
Psychomotor	Increased time needed to complete tasks while using a smartphone	[ES1][ES4][ES6]
	Reduced motor skills	[ES2][ES6]
	Slower movements and reflexes	[ES2][ES5][ES8]
	Rheumatoid arthritis	[ES2][ES8]
	Tighter muscles and joints, and decreased balance	[ES2]
	Difficulty entering data with the finger	[ES3]
	Trembling hands	[ES5][ES8]
	Loss of muscle mass and increased variability in fine motor movements	[ES6]
	Decreased grip strength	[ES6]
	Decreased hand-eye coordination	[ES6][ES7]
	Less precise control of movement	[ES6]
	Decreased retention of hand movement	[ES6]
Cognitive	Reduced memory capacity	[ES1][ES6]
	Difficulty remembering actions and information	[ES2][ES3][ES5][ES6]
	Decreased ability to work and concentrate	[ES2][ES8]
	Requires more time to complete a task	[ES2]
	Difficulty with large amounts of information	[ES2]
	Lack of experience and knowledge of terms	[ES2]
	Confusion	[ES4]
	Difficulty learning new skills	[ES2][ES6][ES8]
	Decreased reasoning, semantic fluency, and spatial cognition	[ES6]
	Decreased cognitive function	[ES7]
	Decreased processing speed	[ES7]
Visual	Reduced visual capacity	[ES1][ES2]
	Lighting issues	[ES1][ES2][ES8]
	Reduced focusing ability	[ES1][ES2]
	Reduced contrast sensitivity	[ES2][ES6]
	Macular degeneration and bifocal lenses	[ES2][ES8]
	Eye strain	[ES2]
	Poor reading visibility	[ES4]

(*continued*)

Table 3. (*continued*)

Category	Identified Characteristic	Secondary Study
	Reduced visual acuity	[ES6]
	Reduced color vision	[ES6]
	Decreased adaptation to darkness	[ES6]
	Reduced visual function	[ES6]
Auditory	Deteriorated hearing ability	[ES2][ES6][ES8]
	Increased likelihood of hearing loss	[ES2][ES8]
Motivational	Frustration and desire to stop using apps due to not identifying the benefit of using them	[ES6][ES8]
	Confidence in one′s own ability	[ES8]
Psychosocial	Different lifestyle goals	[ES6]
	Previous experience with technology	[ES6]
	Slow adoption of technology	[ES6]
Economic	Older adults are more concerned about the costs of devices and internet connections	[ES4]

5 Threats to Validity

In our study, we identified three threats that could affect the validity of our work. First, we only considered three databases for the search process, which means we might have overlooked relevant studies in other databases, thus limiting the number of studies to analyze. Second, analysis errors may be made by the authors of the selected secondary studies, which would impact the information extracted from each of these studies and consequently affect the consistency of the results reported in this work. Finally, the vocabulary used in the search string could have limited the search for studies. However, to obtain this, a control group was used, which consisted of five studies relevant to the research topic, allowing us to understand the vocabulary and identify the most common terms to use in the tested search strings.

6 Conclusions and Future Work

This work reports a tertiary study primarily conducted to identify the characteristics of older adults that can impact the usability of mobile applications. Between 2018 and 2024, we found eight secondary studies. Below are the conclusions for each research question.

RQ1: What Are the Characteristics of Older Adults That Influence or Impact the Usability of Mobile Applications? Older adults present a variety of factors that cause problems when interacting with the interface of a mobile application. Due to these characteristics, the usability of mobile applications can be affected. We grouped the

factors that influence or impact the usability of mobile applications into seven categories: (i) psychomotor limitation, (ii) cognitive limitation, (iii) visual limitation, (iv) auditory limitation, (v) motivational limitation, (vi) psychosocial limitation, and (vii) economic limitation. The categories that stand out the most are the first three mentioned.

RQ2: What Future Research Directions Are Associated With Usability From the Perspective of Older Adult Users? We identified several lines of work highlighting the importance of considering this age group′s specific characteristics and needs when developing technologies and evaluating their usability. For example, developers should consider recommendations for avoiding problems when interacting with mobile applications caused by the characteristics of older adults. It is also necessary to understand the barriers and opportunities related to access and adoption of technology by this age group. Finally, there is a need to make changes to usability evaluation techniques, as these cannot be applied in their original form with the participation of older adults since this age group has characteristics that create challenges when using these evaluation techniques.

Future work includes designing mobile applications with the participation of older adults, considering their particular characteristics, and involving them in usability evaluations. Additionally, it is necessary to determine the most appropriate design and usability evaluation techniques for older adults, considering the characteristics identified in this tertiary study. Furthermore, usability evaluation techniques that cannot be applied in their original version with the participation of older adults must be adapted.

Acknowledgments. The ANID FONDECYT-Iniciación Project supported this work under Grant 11230496 and project DIUDA-22436 of the Universidad de Atacama.

Appendix A: Secondary Studies

[ES1] Elguera, L., Zapata, C.: Elderly users and their main challenges usability with mobile applications: A systematic review. In: Marcus, A., Wang, W. (eds) Design, User Experience, and Usability. Design Philosophy and Theory. HCII 2019, pp. 423–438. Lecture Notes in Computer Science, vol 11583. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-23570-3_31.

[ES2] Ramdowar, H., Khedo, K. K., Chooramun, N.: A comprehensive review of mobile user interfaces in mHealth applications for elderly and the related ageing barriers. Universal Access in the Inform. Society 23(4), 1613–1629 (2024). https://doi.org/10.1007/s10209-023-01011-z.

[ES3] Sinabell, I., Ammenwerth, E.: Challenges and recommendations for eHealth usability evaluation with elderly users: Systematic review and case study. Universal Access in the Information Society 23, 455–474 (2024). https://doi.org/10.1007/s10209-022-00949-w.

[ES4] Awan, M., Ali, S., Ali, M., Abrar, M. F., Ullah, H., Khan, D.: Usability barriers for elderly users in smartphone app usage: An analytical hierarchical process-based prioritization. Scientific Programming 2021(1), 1–14 (2021). https://doi.org/10.1155/2021/2780257.

[ES5] Chalghoumi, H., Al-Thani, D., Hassan, A., Hammad, S., Othman, A.: Research on older persons' access and use of technology in the arab region: Critical overview and future directions. Applied Sciences 12(14), article 7258 (2022). https://doi.org/10.3390/app12147258.

[ES6] Kurniawan, E.; Bakar N.; Salam S.; Restyandito, R.: Usability evaluation dimensions of mobile health application for elderly: A systematic review. Journal of Theoretical and Applied Information Technology 99(1), 242–256 (2021).

[ES7] Aranha, M., James, K., Deasy, C., Heavin, C. Exploring the barriers and facilitators which influence mHealth adoption among older adults: A literature review. Gerontechnology 20(2), 1–16 (2021). https://doi.org/10.4017/gt.2021.20.2.424.06.

[ES8] Wildenbos, G., Peute, L., Jaspers, M.: Aging barriers influencing mobile health usability for older adults: A literature based framework (MOLD-US). International Journal of Medical Informatics 114, 66–75 (2018). https://doi.org/10.1016/j.ijmedinf.2018.03.012.

References

1. World Health Organization: WHO. Envejecimiento y salud. https://www.who.int/es/news-room/fact-sheets/detail/ageing-and-health. Accessed 07 Jan 2025
2. Yadegari, M., Mohammadi, S., Masoumi, A.H.: Technology adoption: an analysis of the major models and theories. Technol. Anal. Strateg. Manag. 1–15 (2022). https://doi.org/10.1080/09537325.2022.2071255
3. Ramírez-Correa, P., Grandón, E.E., Ramírez-Santana, M., Arenas-Gaitán, J., Rondán-Cataluña, F.J.: Explaining the consumption technology acceptance in the elderly post-pandemic: effort expectancy does not matter. Behav. Sci. **13**(2), 87 (2023). https://doi.org/10.3390/bs13020087
4. Briede-Westermeyer, J.C., Pacheco-Blanco, B., Luzardo-Briceño, M., Pérez- Villalobos, C.: Mobile phone use by the elderly: relationship between usability, social activity, and the environment. Sustainability **12**(7), 2690 (2020). https://doi.org/10.3390/su12072690
5. Chopik, W.J.: The benefits of social technology use among older adults are mediated by reduced loneliness. Cyberpsychol. Behav. Soc. Netw. **19**(9), 551–556 (2016). https://doi.org/10.1089/cyber.2016.0151
6. Navarrete, F.D., Needham, T.T., Ortega, A.M.E., Concha, S.M., Macaya, S.X.: Use of information and communication technologies and emotional well-being in older adults. Gac. Méd. Espirituana **24**(2), 1–14 (2022)
7. Castro, J.W., Madrigal, G.A.: Uso de la tecnología por las personas mayores: un estudio exploratorio. In: N. Callaos, J. Horne, E. F. Ruiz-Ledesma, B. Sánchez, A. Tremante (eds.) Memorias de la Décima Cuarta Conferencia Iberoamericana de Complejidad, Informática y Cibernética: CICIC 2024, pp. 137–142 (2024). International Institute of Informatics and Cybernetics. https://doi.org/10.54808/CICIC2024.01.137
8. Londoño Aldana, E., Navas Ríos, M.E., Romero González, Z.: El adulto mayor frente al uso de las herramientas tecnológicas en Cartagena de Indias-Colombia. Aglala **10**(2), 99–110 (2019). https://doi.org/10.22519/22157360.1436
9. Slavicek, T., Balata, J., Mikovec, Z.: Designing mobile phone interface for active seniors: user study in Czech Republic. In: 2014 5th IEEE Conference on Cognitive Infocommunications (CogInfoCom'14), pp. 109–114. Vietri sul Mare, Italy (2014). https://doi.org/10.1109/CogInfoCom.2014.7020428

10. Gonçalves, V.P., et al.: Providing adaptive smartphone interfaces targeted at elderly people: an approach that takes into account diversity among the elderly. Univ. Access Inf. Soc. **16**(1), 129–149 (2015). https://doi.org/10.1007/s10209-015-0429-9
11. Ferreira, S.M., Sayago, S., Blat, J.: Older people positive, active and creative ICT use: a study in three countries. In: Sayago, S. (eds) Perspectives on Human-Computer Interaction Research with Older People, pp. 187–197. Human-Computer Interaction Series. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-06076-3_12
12. Castro, J.W., Lamilla, I.A.: Problemas del adulto mayor con el uso de las aplicaciones móviles. In: N. Callaos, J. Horne, E.F. Ruiz-Ledesma, B. Sánchez, A. Tremante (eds.) Memorias de la Décima Cuarta Conferencia Iberoamericana de Complejidad, Informática y Cibernética: CICIC 2024, pp. 131–136. International Institute of Informatics and Cybernetics (2024). https://doi.org/10.54808/CICIC2024.01.131
13. ISO 9241–11:2018: Ergonomics of human-system interaction part 11: Usability: Definitions and concepts (2nd ed.), ISO (2018)
14. Nielsen, J.: Usability engineering. Morgan Kaufmann Publishers Inc. (1993)
15. ISO 9241–210:2010(E): Ergonomics of human-system interaction part 210: Human-centred design for interactive systems. ISO (2010)
16. Hix, D., Hartson, H.R.: Developing user interfaces: ensuring usability through product & process. John Wiley & Sons, Inc (1993)
17. Peute, L.W., et al.: Overcoming challenges to inclusive user-based testing of health information technology with vulnerable older adults: Recommendations from a human factors engineering expert inquiry. Yearb. Med. Inform. **31**(1), 74–81 (2022). https://doi.org/10.1055/s-0042-1742499
18. Franz, R.L., Barbosa, B., Epp, C.D., Baecker, R., Wobbrock, J.O.: Why and how think-alouds with older adults fail: Recommendations from a study and expert interviews. In: Sayago, S. (eds.) Perspectives on Human-Computer Interaction Research with Older People, pp. 217–235. Human-Computer Interaction Series. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-06076-3_14
19. Sayago, S., Blat, J.: An initial usability evaluation of some word-processing functionalities with the elderly. In: Ninth International Conference on Enterprise Information Systems (ICEIS'07), pp. 328–331. Funchal, Madeira, Portugal (2007)
20. Gao, Q., Ebert, D., Chen, X., Ding, Y.: Design of a mobile social community platform for older chinese people in urban areas. Hum. Factors Ergon. Manuf. Serv. Ind. **25**(1), 66–89 (2012). https://doi.org/10.1002/hfm.20523
21. Kascak, L.R., Rebola, C.B., Sanford, J.A.: Integrating universal design (UD) principles and mobile design guidelines to improve design of mobile health applications for older adults. In: 2014 IEEE International Conference on Healthcare Informatics, pp. 343–348. Verona, Italy (2014). https://doi.org/10.1109/ichi.2014.54
22. Azuddin, M., Malik, S.A., Abdullah, L.M., Mahmud, M.: Older people and their use of mobile devices: Issues, purpose and context. In: 5th International Conference on Information and Communication Technology for the Muslim World (ICT4M'14), Kuching, Malaysia (2014). https://doi.org/10.1109/ICT4M.2014.7020610
23. Isaković, M., Sedlar, U., Volk, M., Bešter, J.: Usability pitfalls of diabetes mHealth apps for the elderly. J. Diab. Res. **2016**, 1604609 (2016). https://doi.org/10.1155/2016/1604609
24. Harrington, C.N., Ruzic, L., Sanford, J.A.: Universally accessible mHealth apps for older adults: towards increasing adoption and sustained engagement. In: Antona, M., Stephanidis, C. (eds.) Universal Access in Human-Computer Interaction. Human and Technological Environments. UAHCI 2017, pp. 3–12. Lecture Notes in Computer Science, vol 10279. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58700-4_1

25. Cáliz, D., Alamán, X.: Usability evaluation method for mobile applications for the elderly: A methodological proposal. In: Pecchia, L., Chen, L.L., Nugent, C., Bravo, J. (eds.) Ambient Assisted Living and Daily Activities. IWAAL 2014, vol. 8868, pp. 252–260. Lecture Notes in Comp. Science. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-13105-4_37
26. Elguera, L., Zapata, C.: Elderly users and their main challenges usability with mobile applications: a systematic review. In: Marcus, A., Wang, W. (eds.) Design, User Experience, and Usability. Design Philosophy and Theory. HCII 2019, vol. 11583, pp. 423–438. Lecture Notes in Comp. Science. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-23570-3_31
27. Kurniawan, E., Bakar, N., Salam, S., Restyandito, R.: Usability evaluation dimensions of mobile health application for elderly: a systematic review. J. Theor. Appl. Inf. Technol. **99**, 242–256 (2021)
28. Wildenbos, G., Peute, L., Jaspers, M.: Aging barriers influencing mobile health usability for older adults: a literature based framework (MOLD-US). Int. J. Med. Inf. **114**, 66–75 (2018). https://doi.org/10.1016/j.ijmedinf.2018.03.012
29. Kitchenham, B., Charters, S.: Guidelines for performing systematic literature reviews in software engineering version 2.3. Technical Report S.o.C.S.a.M. Software Engineering Group, Keele University and Department of Comp. Science University of Durham (2007)
30. Zhang, H., Babar, M.A., Tell, P.: Identifying relevant studies in software engineering. Inf. Softw. Technol. **53**(6), 625–637 (2010). https://doi.org/10.1016/j.infsof.2010.12.010
31. Aranha, M., James, K., Deasy, C., Heavin, C.: Exploring the barriers and facilitators which influence mHealth adoption among older adults: a literature review. Gerontechnology **20**(2), 1–16 (2021). https://doi.org/10.4017/gt.2021.20.2.424.06
32. Sinabell, I., Ammenwerth, E.: Challenges and recommendations for eHealth usability evaluation with elderly users: systematic review and case study. Univ. Access Inf. Soc. **23**, 455–547 (2024). https://doi.org/10.1007/s10209-022-00949-w
33. Ramdowar, H., Khedo, K.K., Chooramun, N.: A comprehensive review of mobile user interfaces in mHealth applications for elderly and the related ageing barriers. Univ. Access Inf. Soc. **23**, 1613–1629 (2024). https://doi.org/10.1007/s10209-023-01011-z
34. Napoleao, B.M., et al.: Establishing a search string to detect secondary studies in software engineering. In: 47th Euromicro Conference on Software Engineering and Advanced Applications (SEAA'21), pp. 9–16. Palermo, Italy (2021). https://doi.org/10.1109/SEAA53835.2021.00010

Generating Product Descriptions Using Customer Reviews on E-Commerce Sites

Aina Ishikawa[1(✉)], Joshujio Takanami[2], and Kohei Otake[3]

[1] Graduate School of Information and Telecommunication Engineering, Tokai University, 2-3-23, Takanawa, Minato-Ku, Tokyo 108-8619, Japan
4cjnm002@tokai.ac.jp

[2] School of Information and Telecommunication Engineering, Tokai University, 2-3-23, Takanawa, Minato-Ku, Tokyo 108-8619, Japan
1cjm1204@mail.u-tokai.ac.ip

[3] School of Economics, Sophia University, 7-1 Kioi-Cho, Chiyoda-Ku, Tokyo 102-8554, Japan
k-otake@sophia.ac.jp

Abstract. In recent years, the scale of the electronic commerce (EC) market has continued to expand along with the spread of the Internet, both for business and personal transactions. Many consumers use EC sites, and it can be said to have penetrated their daily lives. Many consumers make purchasing decisions based on reviews written by other users when purchasing on EC sites. Therefore, reviews are considered to contain a lot of important information indicating the characteristics of products from the viewpoint of users who have purchased and used the products. Until now, many experiments have been conducted on the effects of reviews on consumer behavior and methods for extracting important information from reviews. However, there is room for debate regarding the utilization and validation of the usefulness of information representing product characteristics obtained from reviews. In this study, we propose a method to create product descriptions using information extracted from product review data and to verify the usefulness of such descriptions. Specifically, we identify important words that are considered to represent the characteristics of products by analyzing the sentiment of the reviews and selecting feature words based on the TF-IDF method. We also verify the effectiveness of the feature words by using a screen-based eye tracking device to create product descriptions. The results showed that there was no statistically significant difference in the median gaze duration between the product descriptions created using the feature words obtained from the reviews and the conventional product descriptions.

Keywords: Product Descriptions on E-commerce Sites · Sentiment Analysis · BERT

1 Introduction

In recent years, the Internet penetration rate has dramatically increased. According to the "Results of the 2023 Telecommunications Usage Trends Survey" published by the Ministry of Internal Affairs and Communications, the percentage of individual Internet users is 86.2%, and by age group, the percentage exceeds 90% for each age group

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 57–69, 2025.
https://doi.org/10.1007/978-3-031-93539-8_4

from 13 to 69[1]. The EC (Electronic Commerce) market has been expanding remarkably along with the diffusion of the Internet, and EC is a form of transaction in which goods and services are sold using the Internet communication network, and EC transaction types include BtoB (Business to Business), business-to-business transactions, and business-to-consumer transactions. There are three types of EC transactions: BtoB, BtoC (Business to Consumer), and CtoC (Consumer to Consumer). According to the "FY2023 Market Research Report on Electronic Commerce" by the Ministry of Economy, Trade and Industry (METI), the BtoC-EC market size in the field of goods sales has been growing steadily, and the EC conversion rate is increasing every year [2]. The market size increased by 676.3 billion yen from 13,999.7 billion yen in 2022 to 14,676.0 billion yen, with an increase rate of 4.83% and an EC conversion rate of 9.38%, an increase of 0.25 points from 2022(Fig. 1).

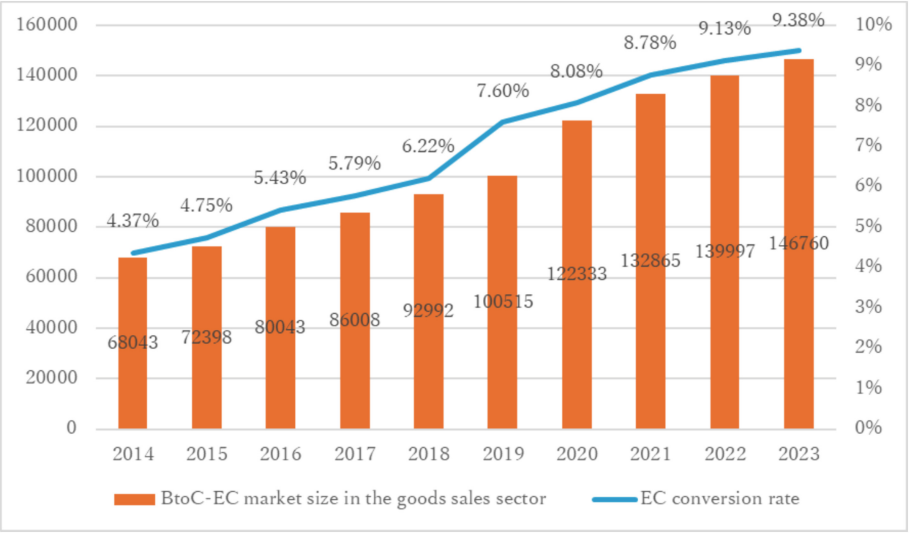


Fig. 1. BtoC-EC market size and EC conversion rate in the goods sales field over time [2].

Furthermore, according to the "EC Consumer Survey 2023" conducted by HAKUHOUDO, which asked about the actual status of EC use and lifestyle awareness among consumers, the EC use rate within one year in 2023 was 83.6%, a high level for all age groups, confirming that EC use has become a common purchasing method among the public [3]. This confirms that the use of EC has penetrated the public as a common means of purchase.

2 Previous Research

With the rapid expansion of the EC market, many studies have been conducted on the effects of product evaluations, word-of-mouth, and reviews on consumer behavior. Sasaki compared the effectiveness of evaluation sites in purchasing products and the

effectiveness of evaluation sites, magazines, and word of mouth among users of evaluation sites [4]. The effectiveness of evaluation websites, magazines, and word of mouth, which are noted for their influence on consumers′ purchasing behavior, were compared through two questionnaires to users of evaluation websites. They found that users of e-commerce websites were more likely to use them. In addition, Eguchi et al. conducted a study on the influence of user reviews on purchasing behavior on e-commerce sites [5]. They revealed that the number of user reviews that highly evaluate a product influences sales, and that the existence of user reviews that are positive about a product and that are referred to by many users positively influences purchasing behavior. On the other hand, some studies have been conducted to extract and utilize important information from review sentences. Yamashita et al. conducted a study on a product review analysis model that extracts features using BERT [6]. They extracted reviews of interest from a large amount of inconsistent product review data using BERT in combination with SVM (Support Vector Machine) and OCSVM (One-Class SVM). They showed that this method can effectively detect "unexpected" and "idiosyncratic" reviews that do not fit the existing labels. In addition, Kamiya et al. proposed and evaluated a summarization method focusing on important words in reviews using deep learning [7]. They introduced an attention mechanism into the summary model generation of the deep learning model to enable summary generation based on important words. A prototype supporting the proposed method was implemented using TensorFlow, and the prototype was adapted to Amazon review data. Evaluation results showed that the Rouge value of the approach can be expected to generate summaries that include important content, independent of the amount of text. Furthermore, Maeda and Yoshioka conducted an experiment using an eye camera to clarify the occurrence of information overload and its effect on purchase decisions for reviews on the Web [8]. In this study, they measured the gaze of experimental participants while browsing a fictitious purchasing site under six conditions combining the quality (impressions and details) and quantity (small, medium, and large) of word-of-mouth. Results showed that when the quantity of word-of-mouth was high, information overload occurred, resulting in a decrease in purchase intention. On the other hand, an increase in the total number of gaze duration was confirmed to promote effective acquisition of information and to increase the willingness to purchase. Thus, eye camera gaze measurement is an effective method for clarifying the relationship between the consumer′s information processing process and purchase decision making.

These studies suggest that consumer evaluations such as user reviews on evaluation sites and e-commerce sites contain useful information and propose a method to extract useful information from the reviews. In addition, there is a suggestion that an approach using gaze data is effective in clarifying the relationship between the consumer′s information processing process and purchase decision making. However, when considering the use of information on reviews in practice, it is necessary to consider how to extract information using a simple approach and how to utilize the information. In addition, by verifying the usefulness of the extracted information in specific situations where it is utilized, it is thought that useful knowledge can be obtained in product marketing and advertising strategies.

3 Purpose of this Study

From previous research, we have learned that user reviews on evaluation sites and e-commerce sites contain important information, and that using gazing data is an effective means of clarifying consumers′ reactions to the information in reviews. In this study, we propose a method for creating product descriptions using information extracted from product review data. Specifically, we identify important words that are expected to express the characteristics of a product through sentiment analysis of reviews and feature word selection based on the TF-IDF method. We also verify the effectiveness of the method using a screen-based eye-tracking device.

4 Dataset

In this study, we used review data on notebook PCs sold by the three companies on a price comparison site [9] and a Japanese subset of the "tyqiangz/multilingual-sentiments" [10] data set provided on Hugging Face. The notebook PCs from the three companies are mainly used for office work and daily use at home, and their price range is from the low 60,000 yen to the high 70,000 yen. The basic specifications are: 2GHz/6-core CPU, 16GB memory capacity, and 512GB storage capacity.

First, we obtained 62 reviews of product A, which is the main subject of analysis in this study. Subsequently, we obtained reviews for Product B and Product C, two products in the same price range as Product A but made by other companies, respectively. We obtained a total of 97 reviews on PCs from the three companies, all of which were included in the analysis. Table 1 below shows the details of the data obtained.

Table 1. Details of Target Data.

Type of product	Number of cases	Period during which the review was posted
A	62	2023/6/23 - 2024/12/5
B	10	2023/09/11 - 2024/11/27
C	25	2024/6/9 - 2024/12/4

We also used the Japanese subset of the "tyqiangz/multilingual-sentiments" [10] dataset provided on Hugging Face for fine-tuning the sentiment analysis model. This dataset is a dataset labeled with sentiment polarity for product reviews on the e-commerce site amazon. The details of the data used are described in the BERT model of sentiment analysis below.

5 Linguistic Processing Analysis on Review Data

First, we conducted sentiment analysis on the acquired reviews and classified them into two categories: positive and negative. Next, we calculated TF-IDF values for words that appeared in both positive and negative reviews to extract words that were considered important in the review of product A.

5.1 Emotion Analysis with BERT Model

First, to extract review sentences with positive content from the acquired reviews, we performed sentiment analysis using BERT and classified them into two categories: reviews with positive content and reviews with negative content.

BERT is a neural language model published by Google in 2018 and is known as a model that shows high performance on a variety of language tasks; in BERT, there are two learning processes: pre-training and fine tuning. In this study, we used "tohoku-nlp/bert-base-japanese-whole-word-masking," a pre-trained Japanese model published by Inui Laboratory at Tohoku University through Hugging Face [11]. First, we performed fine tuning to create a model for the review sentence classification task. For the fine tuning, we used the Japanese subset of the "tyqiangz/multilingual-sentiments" [10] dataset provided on Hugging Face. This dataset consists of train (training data), validation (verification data), and test (validation data), with 120,000 train data and 3,000 each of validation and test (verification data) provided. In this study, we only used 5,000 sampled cases of train data due to the small number of review data we used. We balanced training and validation losses and performed fine tuning with the number of epochs set to 3.

We then used a fine-tuned sentiment analysis model to classify the reviews on the three PCs. Table 2 below shows the results of the classification based on the sentiment analysis.

Table 2. Classification Results of Reviews Based on Sentiment Analysis.

Type of product	Positive	Negative
A	38	24
B	8	2
C	11	14

Table 3 shows the average rating of the actual reviews for each product and the percentage of all reviews with a rating of 5 or 4 out of 5. Compared to the results of the sentiment analysis, PCs with more positive reviews tended to receive higher ratings, while those with more negative reviews received lower ratings. This result does not differ significantly from the results of the sentiment analysis.

Table 3. Actual Review Rating Values and Percentage of High Ratings.

Type of product	Average Rating	percentages of reviews scoring 5 or 4 out of 5
A	4.32	91%
B	3.89	76%
C	4.17	88%

5.2 Extraction of Important Words and Creation of Product Descriptions Using TF-IDF

Next, we calculated TF-IDF values for words that appeared in both positive and negative reviews to extract words that were considered important in the reviews of product A.

First, we performed morphological analysis on the 57 reviews classified as positive, and extracted only nouns, adjectives, and adverbs whose meanings were comprehensible as stand-alone words. In this study, we considered adverbs to be important information because they are often used as modifiers to PC movements. In addition, we deleted words such as "purchase," "use," and "ordinary," which were not considered to represent the characteristics of the product when creating the product description. Next, we extracted feature words related to the review of Product A by calculating TF-IDF values for the words that appeared in the review of Product A. In this study, we used the TF-IDF values of the words that appeared in the review of Product A to extract feature words related to the review. For calculating TF-IDFs, we have calculated TFs for the 38 positively classified reviews of product A as one document and IDFs for the 57 positively classified reviews of the three products as one document. Table 4 below shows the top 15 words with the highest TF-IDF values in the review documents for Product A.

Table 4. Top 15 Words with High TF-IDF Value for Positive Reviews.

Word	TF-IDF value
USB (USB)	48.268
デザイン(design)	30.000
拡張(expansion)	29.000
コストパフォーマンス(cost performance)	28.000
速度(speed)	28.000
[Microprocessor brand name]	27.282
SSD (SSD)	27.282
キーボード(keyboard)	27.000
処理(prosessing)	27.000
電源(power supply)	26.704
性能(perfofmance)	26.000
HP (HP)	23.893
Dell (Dell)	23.085
グラフィック(graphics)	23.000
問題(issue)	23.000

These words are important words that represent the characteristics of product A. However, if the same feature words are extracted from the negatively categorized reviews of Product A, and words common to those in Table 4 are extracted, those words are not necessarily used favorably by the consumers. To create product descriptions that are attractive to consumers, we need to use featured words that appear only in positive reviews. Therefore, we calculated TF-IDF values for negative reviews using the same

method. Table 5 below lists the top 15 words with the highest TF-IDF values in negative review documents.

Table 5. Top 15 words with high TF-IDF value for negative reviews.

Word	TF-IDF value
USB(USB)	31.479
性能(perfofmance)	28.000
価格(price)	21.081
デザイン(design)	21.000
コストパフォーマンス(cost performance)	20.000
処理(processing)	20.000
拡張(expansion)	19.676
快適(comfort)	15.460
端子(terminals)	15.460
グラフィック(graphics)	15.000
インチ(inch)	14.054
動画(video)	14.054
バッテリ(battery)	14.000
作業(work)	14.000
速度(speed)	14.000

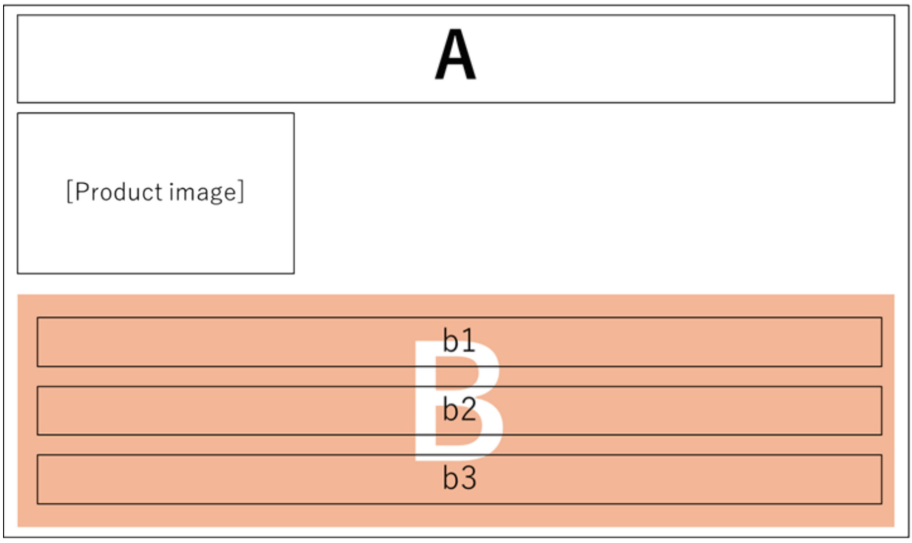


Fig. 2. Scenario Product Page Model.

Table 6. Description of the Product Per Condition (Area B).

Target Condition	Area		Description
Condition 1 and Condition 2	A		【メーカー名】【MPUブランド名】・16GBメモリ・512GB SSD・フルHD・IPSパネル搭載 価格.com限定モデル [ナチュラルシルバー] ([Manufacturer] [Microprocessor brand name], 16GB memory, 512GB SSD, Full HD, IPS panel, Price.com limited model [Natural Silver].)
Condition 1	B	b1	【マイクロプロセッサブランド名】は、高い省電力性能を備えた最新のモバイルプロセッサで、スムーズなマルチタスクと優れたグラフィック性能を提供。(The [microprocessor brand name] is the latest mobile processor with high power-saving performance, offering smooth multitasking and excellent graphics performance.)
		b2	512GB SSDは、高速なデータ転送速度と十分な保存容量を備えたストレージで、快適なシステム起動とアプリケーション動作を実現。(512GB SDD provides fast transfer speeds and ample storage capacity for comfortable system startup and application operation.
		b3	快適作業を支えるテンキー付きキーボード。(A keyboard with numeric keypad for comfortable working.)
Condition 2	B	b1	15.6型フルHD（1920×1080）ディスプレイを採用したノートPC．IPSなので、斜めからも見やすくより自然な映像を表示。(The 15.6-inch full HD (1920 x 1080) display on this notebook PC is IPS, making it easy to view images even from an angle and providing a more natural look.)
		b2	（バッテリー駆動時間は最大10時間，45分で50%充電できる「ファストチャージ（急速充電機能）」を備える。(The battery life is up to 10 hours and can be charged to 50% in 45 minutes with “Fast Charge.)
		b3	USBポートを介さなくても使用できるBluetoothマウスが付属．指紋認証でのログインにも対応し、マスクをしていても素早くログイン可能。(A Bluetooth mouse is included, which can be used without a USB port. Fingerprint authentication is also supported for quick login even when wearing a mask.)

Based on these results, we established the criteria that the words must not be among the words with high TF-IDF values in the negative content reviews (Table 5) and must have high TF-IDF values in the positive content reviews. Based on this criterion, we selected three words: "(microprocessor brand name)," "SSD," and "keyboard. Table 6 shows the contents of each area of the scenario product page that contains the product description created using the three selected words for Condition 1 and the conventional product description for Condition 2. (Note that the images presented to the subjects include the names of actual manufacturers and brands.) Fig. 2 shows a model of a scenario product page. Area B contains short sentences describing the product in detail from three perspectives. Area A and the product image are common to both condition 1 and Condition 2. Furthermore, we set the area containing the product description as the AOI (Area of Interest) (Fig. 2). (Note that the AOI areas in Fig. 2 are colored for convenience, but the images presented to the subjects are not colored.)

6 Screen-Based Eye Tracking Experiment to Test the Usefulness of Created Product Descriptions

Next, we conducted a screen-based eye tracking experiment using a Tobii eye tracker to test the usefulness of the product descriptions we created.

6.1 Experimental Design

The experiment was conducted at Tokai University on January 6, 2025. The subjects were 8 consenting Tokai University students (3 males, 5 females, age range: early 20s).

In this experiment, we used the Tobii Pro Spark eye tracking device to acquire eye tracking data (Fig. 3). This device is a screen-based eye tracker that can acquire eye tracking data at 60 Hz.

Fig. 3. Tobii Pro Spark [12]

We conducted our experiments following the flow shown in Fig. 4 below. First, we performed a calibration (calibration is the process of acquiring the subject′s eye and geometric features to accurately calculate eye gaze). The first step was to perform a calibration of the subject′s eyes. Next, we measured eye gazing data for the two scenario product pages (Condition 1 and Condition 2). We told the subjects in advance, "Please browse the product pages on the two notebook PCs as you would normally browse an e-commerce site, but please note that the product images are not displayed on both notebook PCs. However, the product images are the same for both." The eye tracker was installed in the eye tracking system. We displayed each scenario for 20 s on the screen of the laptop equipped with the eye tracker. We switched the order of the images displayed for each subject to account for order effects. We also provided a 3-s break by displaying

a white screen between each scenario. The purpose of the break was to reset the subject′s cognition, preventing the effects of the previous scenario from being carried over to the next scenario, and at the same time, to reduce the subject′s cognitive burden and fatigue.

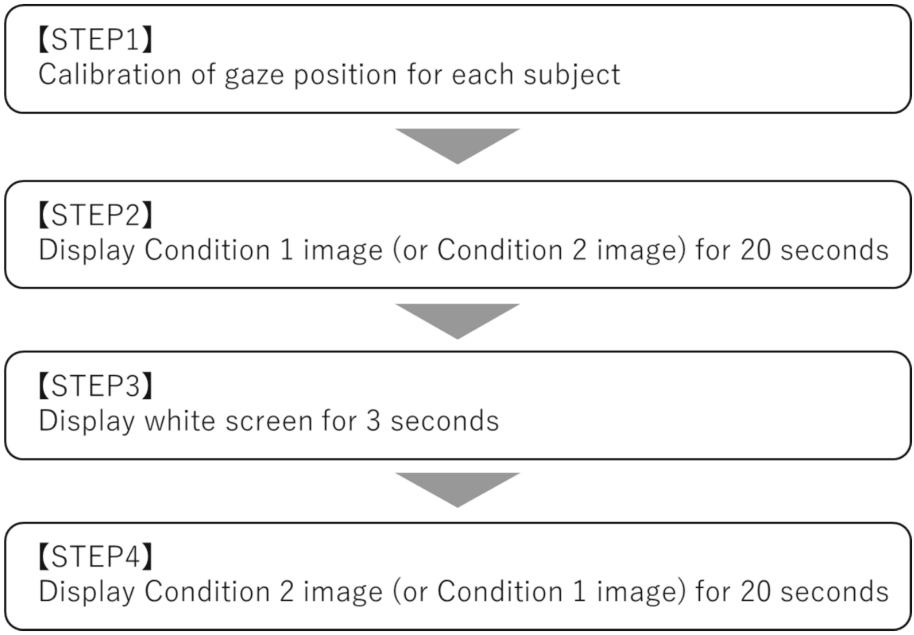


Fig. 4. Experiment Flow.

Next, we performed a heatmap visualization using Tobii pro Lab [13] on a scenario product page containing the created product description (Fig. 5) and a conventional product description (Fig. 6), using eye gazing data from eight subjects.

Figure 5 and Fig. 6 show that the number of areas that were gazed at (red areas) increased in Condition 1 compared to Condition 2. In particular, the user′s eyes tended to focus on specific and numerical words such as “45 min,” “50%,” “512GB,” and “SSD. These words are likely to attract users′ attention as elements that clearly convey the performance and features of the product. This suggests that designing explanatory text that emphasizes concreteness may promote visual attention and contribute to better product understanding.

The gaze duration at the AOI (area B in Fig. 2) was calculated using gazing data from the two scenarios. Gaze duration is calculated as the number of records hit in the AOI / the number of sampling rates (60 Hz for Tobii pro spark). The gaze duration was calculated as the average gaze duration for eight subjects for each scenario was 10.25875 s for condition 1 and 12.0875 s for condition 2. For the differences in gaze duration within region B shown in Fig. 2, we performed a Wilcoxon signed rank test. It is applied when the data cannot be assumed to follow a normal distribution. In this study, the null hypothesis (H_0) is there is no statistically significant difference in gaze duration between the created product descriptions (condition 1) and the traditional product descriptions

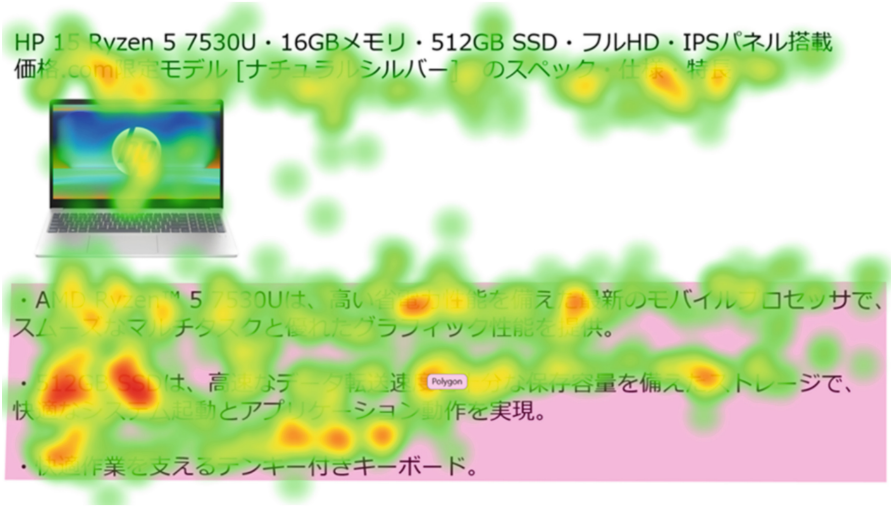


Fig. 5. Heatmap of the Scenario Product Page Including the Created Product Description (Condition 1). (Color figure online)

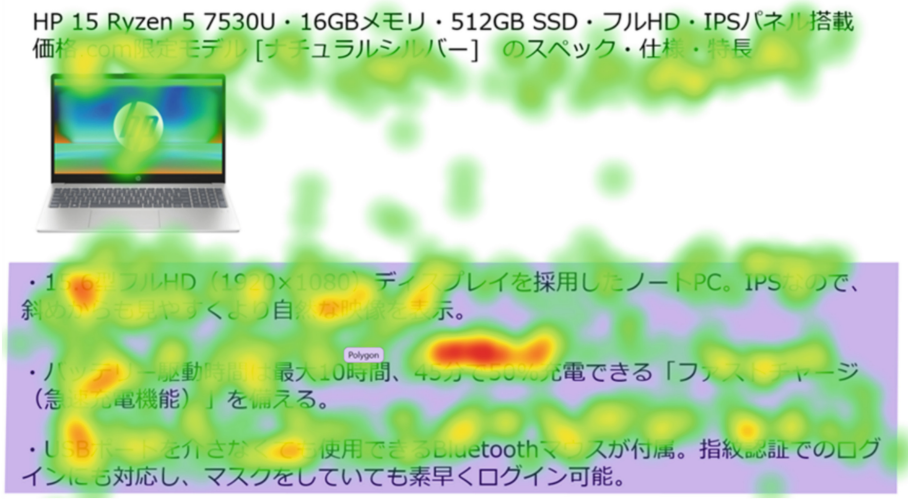


Fig. 6. Heatmap of the Scenario Product Page Including the Conventional Product Description (Condition 2). (Color figure online)

(condition 2)" and the alternative hypothesis (H_1) is "there is a statistically significant difference between the created product descriptions (condition 1) and the traditional product descriptions (condition 2)". The test results showed that the test statistics was 7.0 and the p-value was 0.1484375. At a significant level of 5%, the null hypothesis was not rejected. In other words, it was concluded that there is no significant difference in gaze duration between the product descriptions in Condition 1 and Condition 2. On

the other hand, the effect size (r), at 2.475, is very large, suggesting that there may be substantial differences between conditions. As can be seen from the heatmap in Fig. 5, this may be because the effect size increased not for the entire product description, but for the areas where the gaze duration was particularly long (red colored areas).

7 Conclusion

In this study, we conducted a binary classification of reviews by using a fine-tuned BERT model with a dataset of positively and negatively labeled product review sentences on an e-commerce site. As a result, we were able to confirm the correlation with the evaluation values in actual product reviews. In addition, we were able to extract words that were considered important to consumers by conducting a frequency analysis using TF-IDF for each of the review documents classified as positive and negative. Specifically, words such as "CPU," "graphics," and "design" with positive sentiment and words such as "speed," "model," and "battery" with negative sentiment were extracted. These words indicate what expectations consumers have of the product and in which areas they are dissatisfied. We believe that reflecting the words extracted by this method in product descriptions and advertising copy will be effective in attracting consumers and motivating them to purchase the product. Emphasizing positive words and avoiding negative words may improve consumers′ impression of the product. In addition, by visually highlighting important words in the description, consumers are more likely to pay attention to important information, which is expected to lead to an increase in purchase intention.

Future work include securing the sample size for the experiment and verifying the results by considering other factors such as the amount of text and design. In this experiment, we could not confirm statistically significant differences in gaze duration between product descriptions created using words extracted from review documents and conventional product descriptions. On the other hand, the effect size (r) was very large at 2.475, suggesting that there may be a substantial difference between the conditions. This result suggests that statistical significance could not be detected due to insufficient sample size or other external factors. Therefore, we believe that further validation is needed to increase the sample size and consider the influence of other factors. In this experiment, we set the AOI for the entire product description but did not set detailed AOIs at the word or phrase level. We believe that more precise results can be obtained by setting multiple AOIs for each word or phrase and verifying them. Although further improvement and accuracy enhancement are needed in the future, the approach taken in this study is expected to be an effective method in product marketing and advertising strategies.

Acknowledgements. We would like to express our sincere gratitude to all the subjects who cooperated in this study. This work was supported by JSPS KAKENHI Grant Number 24K05159.

References

1. Ministry of Economy, Trade and Industry: 2023 Market Research Report on Electronic Transactions. https://www.soumu.go.jp/johotsusintokei/statistics/statistics05.html (2024/12/14 author checked)

2. Ministry of Economy, Trade and Industry: FY2023 Market Research Report on Electronic Transactions. https://www.meti.go.jp/policy/it_policy/statistics/outlook/R5tyousahoukokusho.pdf (2025/1/19 author checked)
3. HAKUHODO: Hakuhodo DY Group Shopper Marketing Initiative® HAKUHODO EC + x Hakuhodo Shopping Institute, EC Sei-katsu-sha Survey 2023 (in Japanese). https://www.hakuhodo.co.jp/uploads/2023/06/H20230627-1.pdf (2024/12/14 author checked)
4. Sasaki, Y.: The Effectiveness of Rating Websites in Product Purchasing: A Comparison of the Effectiveness of Rating Websites, Magazines, and Word of Mouth among Users of Rating Websites Information Media Research, vol. 3, no. 1, p. 29–42 (2004) (2025/1/19 author checked)
5. Egawa, Y., Ichifuji, Y., Konno, M.: A study on the influence of user reviews on purchasing behavior in e-commerce sites information processing society of Japan. In: Proceedings of the 73rd National Convention, no. 11, p. 591−592 (2011) (in Japanese)
6. Yamashita, K., Ito, F., Hasumoto, K., Goto, M.: A study on product review analysis model using feature extraction by BERT information processing society of japan. In: Proceedings of the 84th National Convention, no. 1, p. 841−842 (2022) (in Japanese) (2025/1/19 author checked)
7. Kamiya, K., Hara, H., Aoyama, M.: Proposal and evaluation of a summarization method focusing on important words in reviews by deep learning processing society of Japan, no.1, p.463−464 (2020) (in Japanese) (2025/1/19 author checked)
8. Maeda, H., Chisaki, Y.: Effects of quantity and quality of word-of-mouth on the Web on information overload and purchase decision making on the Web. Bull. Kyoto Tachibana Univ. **48**, 277−291 (2022) (2025/1/19 author checked)
9. kakaku.com official website. https://kakaku.com/ (2024/12/14 author checked)
10. Hugging Face: datasets/tyqiangz/multilingual-sentiments. https://huggingface.co/datasets/tyqiangz/multilingual-sentiments/viewer/japanese (2025/1/19 author checked)
11. Hugging Face: tohoku-nlp/bert-base-japanese-whole-word-masking. https://huggingface.co/tohoku-nlp/bert-base-japanese-whole-word-masking (2024/12/14 author checked)
12. tobii: Tobii Pro Lab/Spark new. https://www.tobii.com/ja/company/newsroom/tobii-pro-spark-new (2024/12/14 author checked)
13. Tobii: Tobii Pro Lab eye tracking research software. https://www.tobii.com/ja/products/software/behavior-research-software/tobii-pro-lab (2024/12/14 author checked)

Improving Intention Recognition Efficiency: A Study on Skeletal Data Dimensionality Reduction and Neural Architectures

Luka Orsag[1,2(✉)], Tomislav Stipancic[1], Leon Koren[1], and Matija Zidaric[1]

[1] Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, Zagreb, Croatia
luka.orsag@visagetechnologies.com
[2] Visage Technologies AB, Linköping, Sweden

Abstract. Understanding human intent is crucial for human-computer interaction (HCI) and human-robot collaboration (HRC) in industrial environments. However, existing intention recognition models suffer from poor generalization, high computational costs, and latency issues, particularly when processing high-dimensional skeletal motion data. This study explores dimensionality reduction techniques, including Principal Component Analysis (PCA), Kernel PCA (kPCA), and Uniform Manifold Approximation and Projection (UMAP), to optimize feature selection while maintaining relevant motion features.

A multimodal system using RGB, depth, and embedded sensors is developed to classify human intentions. The proposed system integrates neural network-based models to analyze spatial and temporal motion patterns, enabling adaptive robotic behavior through short-term pose prediction, collision sensing, and path correction. The OAK-D AI platform is employed for real-time depth perception and edge AI processing.

While the applied dimensionality reduction techniques lead to reduced model accuracy, they significantly enhance computational efficiency, lower processing latency, and reduce power consumption. These trade-offs make the system more suitable for real-time, resource-constrained industrial applications, ensuring faster but slightly less accurate decision-making in AI-driven industrial automation and workplace safety monitoring.

Keywords: Intention recognition · Dimensionality reduction · Skeletal motion analysis

1 Introduction

Intention recognition is a foundational aspect of human-computer interaction (HCI), enabling systems to anticipate and respond to human actions in a seamless and intuitive manner. It has applications in robotics, wearable technologies, gesture-controlled interfaces, and assistive systems. By understanding human intent, machines can improve interaction quality, safety, and efficiency [1, 3]. However, achieving reliable intention

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 70–86, 2025.
https://doi.org/10.1007/978-3-031-93539-8_5

recognition often requires analyzing skeletal motion data, which is inherently high-dimensional. Each motion capture dataset, such as those based on body joint coordinates, can include hundreds of features per sample [2, 4].

Camera-based systems like OpenPose extract skeletal data by identifying body joints in 2D or 3D spaces [2]. While accurate, these systems generate large datasets with highly redundant information, leading to inefficiencies when used directly in neural networks [3]. High-dimensional data also complicates the integration of other modalities, such as head tracking or body-worn sensor inputs, which are often essential for robust intention recognition in dynamic, real-world scenarios [5].

These computational and scalability challenges make it difficult to deploy intention recognition systems in real-time and resource-constrained environments like wearable devices or robotics. Consequently, developing methods to streamline skeletal data while retaining its most critical features is essential for advancing HCI systems [4, 5].

Despite significant advancements in intention recognition, current methods often rely heavily on high-dimensional skeletal data generated by camera-based systems like OpenPose [2]. While these systems provide detailed representations of human motion, they introduce several challenges. High-dimensional datasets increase computational overhead, require significant storage resources, and make real-time processing in resource-constrained environments such as robotics or wearable devices infeasible [3, 6].

Moreover, the reliance on visual modalities limits the adaptability of these systems. In dynamic environments where lighting, occlusions, or background clutter may impact camera performance, the models become less reliable [1, 4]. This limitation also complicates the integration of multimodal data sources, such as head tracking or body-worn sensors, which are critical for capturing intentions beyond what visual data alone can provide [5].

While there has been progress in reducing the dimensionality of skeletal data, many studies focus on automated methods such as deep learning feature extraction, which may not fully account for task-specific nuances. Manual or semi-automated approaches, informed by domain expertise, are underexplored in this context, leaving a gap in leveraging efficient feature selection to optimize skeletal datasets for specific applications like intention recognition [3, 7]. Additionally, the challenges associated with high-dimensional data, such as the risk of overfitting and the need for large sample sizes to achieve reliable models, further complicate the development of efficient intention recognition systems [6, 7]. Addressing this gap requires a comprehensive investigation into dimensionality reduction methods that balance computational efficiency, multimodal adaptability, and the preservation of critical features necessary for robust intention recognition.

Dimensionality reduction is a crucial technique for handling high-dimensional skeletal data in industrial intention recognition. By reducing the number of features while retaining critical information, it enhances computational efficiency, reduces storage demands, and improves model interpretability. Several techniques, including Feature Selection, Principal Component Analysis (PCA), Kernel PCA, and Uniform Manifold Approximation and Projection (UMAP), are widely applied in industrial settings for optimizing real-time performance.

PCA, a widely used linear technique, transforms data into uncorrelated principal components that capture the maximum variance within skeletal motion data. This method has been effectively used in industrial imaging and defect detection, where PCA-based dimensionality reduction improves segmentation accuracy and processing speed in large-scale image datasets [8].

For non-linear data, Kernel PCA extends traditional PCA by enabling higher-order transformations that preserve complex relationships within skeletal motion datasets. This method has been successfully applied in robotic intention recognition and industrial automation, where Kernel PCA improves feature extraction from high-dimensional sensor data, enabling better classification and decision-making [9].

Beyond PCA, UMAP offers an alternative approach for preserving both local and global structures in high-dimensional data. Unlike PCA, which primarily focuses on variance, UMAP constructs a graph-based representation of the data, making it highly suitable for real-time intention recognition in industrial settings. Studies have shown that UMAP enhances classification accuracy for damage detection in wind turbines [10] and biomedical industrial applications [11]. Furthermore, UMAP has proven to be particularly effective in human-robot interaction, where it improves gesture-based intention recognition by reducing redundant skeletal motion features while retaining key information for classification [12].

These dimensionality reduction techniques not only optimize intention recognition models but also enable the integration of multimodal industrial data sources, such as wearable sensors or skeletal tracking systems, without overburdening computational resources. By selecting the most relevant features while discarding redundant ones, PCA, Kernel PCA, and UMAP play a fundamental role in enhancing machine learning-driven decision-making in industrial applications.

The analysis of human actions and intentions has been extensively studied, with deep learning models such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks proving to be highly effective. These methods allow for the extraction of spatiotemporal features from skeletal data, improving action recognition accuracy. Graph Convolutional Networks (GCNs) further enhance performance by capturing spatial and temporal dependencies in skeletal movement data, making them particularly effective for understanding human motion in industrial environments [16, 17].

Transfer learning has also emerged as a key technique, enabling models to generalize to new domains with minimal annotated data. This reduces the burden of extensive data labeling while maintaining high recognition accuracy [17]. Hierarchical models, which integrate different neural network architectures, have been developed to capture complex relationships between human activities, offering a more robust understanding of worker behavior [18].

However, challenges remain in action recognition, particularly in industrial settings where diverse environments and complex interactions create difficulties in real-time recognition. A significant focus has been placed on improving system responsiveness while reducing computational delays, ensuring that real-time applications remain efficient and accurate [19]. Additionally, data quality and annotation complexity

remain persistent issues, requiring innovative solutions to streamline data collection and preprocessing [20].

A promising approach to enhancing model robustness is the integration of multi-sensor and multimodal data, such as head tracking and wearable sensors, allowing for more accurate and adaptive action recognition. Extracting relevant skeletal features from raw data remains a critical step, ensuring that models remain scalable and computationally feasible. Moreover, domain adaptation and advanced transfer learning techniques offer significant improvements in model generalization, particularly when applied to industrial settings.

Building upon this foundation, authors revisited their LSTM-based activity and intention recognition framework, further refining their approach to human-robot collaboration (HRC) by incorporating head pose estimation and context-aware learning [16, 17]. Their studies highlight the importance of spatiotemporal modeling, allowing intelligent systems to better predict human intentions in real-time environments.

Recent advancements have introduced Transformer-based architectures, significantly improving sequential data processing. Unlike traditional RNNs or LSTMs, Transformers leverage self-attention mechanisms, enabling them to efficiently capture long-range dependencies within skeletal motion sequences. The work by Vaswani et al. (2017) introduced the "Attention is All You Need" framework, demonstrating how self-attention mechanisms outperform traditional sequence-processing architectures [22].

Expanding upon this, the Action Transformer model has been proposed to refine human action recognition by emphasizing spatial and temporal attention over key skeletal movements. This technique enables fine-grained activity analysis, particularly in industrial scenarios where rapid and accurate decision-making is crucial [23]. Studies have shown that Transformer-based models achieve higher accuracy and robustness in real-world industrial applications, outperforming traditional LSTM and CNN models in complex motion prediction tasks [24].

By combining Transformer networks with GCNs and LSTMs, future models are expected to further enhance real-time action recognition, reducing computational costs while improving accuracy and adaptability in industrial human-robot collaboration.

2 System Design

The proposed system presented in Fig. 1, leverages multimodal signal acquisition, feature extraction, and neural network-based classification to enhance human-robot collaboration (HRC). Using RGB, depth, and embedded vision sensors, it captures spatio-temporal motion data, which undergoes feature selection through facial and skeletal tracking. A neural network classifies human intentions and activities by analyzing spatial and temporal motion patterns, enabling real-time recognition and prediction. This facilitates adaptive robotic behavior, supporting short-term pose prediction, collision sensing, and path correction. By integrating AI-driven motion classification, the system ensures safe and efficient HRC through proactive decision-making. Additionally, the use of InHARD dataset [25] maximizes the applicability of the system in industrial scenarios.

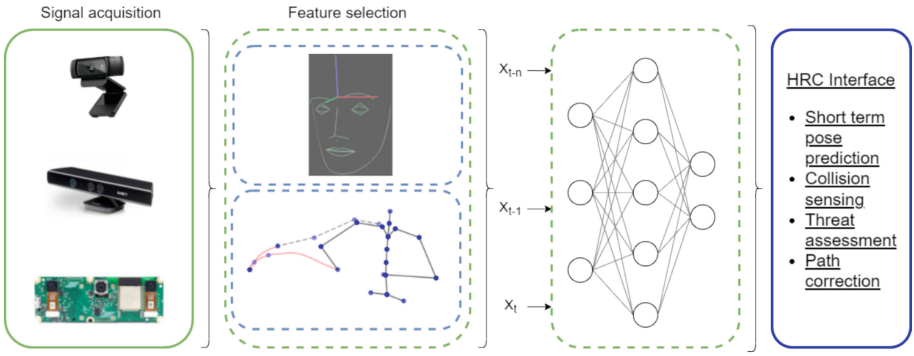


Fig. 1. Overall system design.

2.1 OAK-D AI Platform with Depth Perception

OAK-D (OpenCV AI Kit with Depth) is a state-of-the-art embedded spatial AI device designed for real-time computer vision applications. It integrates a stereo depth camera with an RGB camera and an Intel Myriad X VPU (Vision Processing Unit), enabling efficient edge computing for AI tasks.

The stereo depth system provides real-time 3D perception, allowing applications such as robotics, autonomous navigation, and augmented reality. The onboard Neural Processing Engine (NPE) facilitates deep learning inference, supporting frameworks like TensorFlow, PyTorch, and OpenVINO. Unlike conventional cameras, OAK-D performs depth estimation and object detection directly on the device, reducing reliance on external computation.

We deploy OAK-D because of it's high configurability and ability to run various neural network architectures along with it's hardware advantages (RGB + Depth cameras).

2.2 Inhard Dataset Dimensionality Reduction

Principal Component Analysis (PCA) for Joint Selection. Let a skeletal motion sequence be represented as a matrix $X \in \mathbb{R}^{N \times D}$, where N is a total number of frames across all sequences and $D = J \times 3$ is the total number of features, with J being the number of skeletal joints.

The goal of PCA is to transform the data into a lower dimensional space while preserving as much variance as possible. This is achieved by computing the eigen decomposition of the covariance matrix.

Since PCA is sensitive to scale, we first standardize the dataset by centering and scaling each feature:

$$\tilde{X} = \frac{X - \mu}{\sigma} \tag{1}$$

where μ is the mean vector and σ is the standard deviation across all frames.

The covariance matrix $C \in \mathbb{R}^{D\times D}$ is computed as:

$$C = \frac{1}{N-1}\tilde{X}^T\tilde{X} \tag{2}$$

The eigenvalue decomposition is then performed on the covariance matrix:

$$Cv_i = \lambda_i v_i, i = 1,2,\ldots,D$$

where v_i are the eigenvectors and the λ_i are the eigenvalues, representing the amount of variance explained by each component.

Since each joint contributes with three coordinates to the feature space, it's total contribution to the variance can be computed as:

$$S_j = \sum\nolimits_{k=1}^{K} \sum\nolimits_{d\in\{x,y,z\}} \left| v_{kd}^{(j)} \right| \tag{3}$$

where S_j is the total contribution score for joint j, $v_{kd}^{(j)}$ represents the absolute loading of the k-th principial component and K is the number of principial components considered. By ranking the joints based on the contribution we select top $J\prime \ll J$ most important joints for further analysis. Using those selected joints, we construct the feature set $X\prime \in \mathbb{R}^{N\times(J\prime\times 3)}$ for subsequent processing. We perform the procedure on the complete dataset and the dataset split by labels. This PCA-based selection ensures that only the most informative joints are retained.

Kernel Principial Component Analysis (kPCA) for Joint Selection. While Principal Component Analysis (PCA) effectively identifies linearly correlated joints, it fails to capture nonlinear dependencies in skeletal motion data. Kernel PCA (kPCA) extends PCA by mapping the original data into a higher-dimensional feature space where nonlinear structures become linearly separable. This section formulates kPCA for selecting the most informative joints in motion sequences. Starting from the same dataset X as used for the PCA method, the goal of kPCA is to project data into a high-dimensional feature space using a kernel function $\phi : \mathbb{R}^D \to \mathbb{H}$, where $\mathbb{H}$ is a reproducing kernel Hilbert space (RKHS).

Instead of explicitly computing the $\phi(X)$, we define a kernel function $k(X_i, X_j)$ that computes inner products in the feature space:

$$K_{ij} = k\big(X_i, X_j\big) = \langle \phi(X_i), \phi(X_j)\rangle \tag{4}$$

The kernel of choice for this work is the Radial Basis Function (RBF) kernel:

$$k\big(X_i, X_j\big) = \exp\left(\frac{\|X_i - X_j\|^2}{2\sigma^2}\right) \tag{5}$$

Since we operate in a feature space without explicit feature vectors, we center the kernel matrix $K\epsilon\mathbb{R}^{N\times N}$:

$$K' = K - 1_N K - K1_N + 1_N K 1_N \tag{6}$$

where 1_N is an *NxN* matrix with all elements equal to 1/*N*. The eigenvalue decomposition is then solved:

$$K'v_i = \lambda_i v_i \tag{7}$$

Since kernel PCA does not provide explicit feature loadings like PCA, we compute joint contributions using a leave-one-joint-out (LOJO) [14]. For each joint j:

- Remove all features associated with the joint j
- Perform kPCA again and compute total variance retained:

$$V_j = \sum_{i=1}^{K} \lambda_i^{(j)} \tag{8}$$

where $\lambda_i^{(j)}$ are eigenvalues of kPCA performed after removing joint j.

- Define joint contribution score as the drop in variance:

$$S_j = V_{full} - V_j \tag{9}$$

Like the PCA method we rank the joints using the S_j.

Uniform Manifold Approximation and Projection (UMAP) for Joint Selection. While PCA and Kernel PCA provide linear and nonlinear feature selection, respectively, they do not explicitly preserve the local neighborhood structure of skeletal data. The method UMAP addresses this limitation by learning a low dimensional representation that retains both global and local structures. This section formulates UMAP for identifying the most informative joints in skeletal motion sequences. The method constructs a graph-based embedding where data points are connected based on their local neighborhood relationships, forming lower-dimensional representation.

Instead of assuming a linear transformation, UMAP builds a weigted graph $G = (V,E)$ where nodes V represent frames and edges E connect similar frames based on a fuzzy topological structure. The probability between points X_i and X_j is computed using a kernel function:

$$p_{ij} = exp\left(-\frac{\|X_i - X_j\|^2}{\sigma_i}\right) \tag{10}$$

where σ_i is a local connectivity parameter. The global fuzzy representation is:

$$P = 1 - \prod_{i=1}^{N}(1 - p_{ij}) \tag{11}$$

UMAP minimizes the cross-entropy loss between the high-dimensional probability distribution P and a low-dimensional distribution Q in the reduced space:

$$\mathcal{L} = \sum_{(i,j)\in E} P_{ij}\log Q_{ij} + (1 - P_{ij})\log(1 - Q_{ij}) \tag{12}$$

where:

$$Q_{ij} = (1 + a\|Y_i - Y_j\|^b)^{-1} \tag{13}$$

is a parametric function that models neighbor relationships? Parameters a and b control the shape of embeddings.

Since UMAP doesn't provide explicit feature coefficients like PCA, we compute joint contributions using SHAP (Shapley Additive Explanations) [15]. First we fit UMAP on a standardized dataset Y = UMAP(X) where $Y \in \mathbb{R}^{N \times K}$ is the low-dimensional representation and then train SHAP explainer using the learned embeddings:

$$S_j = \frac{1}{N} \sum_{i=1}^{N} \left|\phi_j(X_i)\right| \tag{14}$$

Using S_j we rank and select top J' most important joints.

kPCA and UMAP Complexity Reduction. Due to the high computational load associated with feature importance estimation, we employed several optimization strategies to enhance efficiency. For kPCA, we utilized the Nyström approximation [13], which approximates the kernel matrix K using a subset of data points, reducing computational complexity from $O(N^3)$ to $O(mN^2)$, where $m \ll N$:

$$\tilde{K} = K_{NM} K_{MM}^{-1} K_{MN} \tag{15}$$

where K_{NM} represents the kernel matrix between the full dataset and the sampled subset, and K_{MM} is the kernel matrix of the sampled subset. Additionally, we implemented a batch-wise computation strategy, partitioning data into smaller batches of size BBB and processing each independently:

$$X_{batch} = f(X[i : i + B]) \tag{16}$$

where X_{batch} represents the reduced-dimensional output for a given batch i, preventing memory overflow while preserving accuracy. These optimizations significantly improved processing time and memory efficiency across PCA, Kernel PCA, and UMAP, enabling large-scale joint selection.

Training a Random Forest Classifier (RFC) for Method Evaluation. To validate the effectiveness of different dimensionality reduction techniques for skeleton-based motion classification, we employ a Random Forest Classifier. RFC is a widely used ensemble learning method that combines multiple decision trees to enhance classification robustness and reduce overfitting. Given its ability to handle high-dimensional structured data, it serves as an appropriate baseline for evaluating the impact of feature selection on model performance.

We utilize Scikit-Learn's RandomForestClassifier, which consists of an ensemble of decision trees trained using bootstrapped samples. Given its ability to implicitly handle feature importance, RFC is well-suited for evaluating the discriminative power of selected joints.

The dataset is split into training (80%) and testing (20%) subsets using stratified sampling:

$$(X_{train}, y_{train}), (X_{test}, y_{test}) = train_test_split(X, y, test_size = 0.2) \tag{17}$$

where stratification ensures class balance between splits. After fitting, predictions are obtained:

$$\hat{y} = RFC.predict(X_{test}) \tag{18}$$

2.3 Intention Classification Model Architectures

This section presents four different deep learning models applied to intention recognition: LSTM networks, Transformers, CNN-based models and a Hybrid TCN-Transformer model. We also analyze their compatibility with edge deployment on OAK-D.

Long Short-Term Memory Model. LSTM networks are designed to handle sequential dependencies in time-series data. Unlike traditional recurrent networks, LSTMs use memory cells to retain long-term information. Each LSTM cell consists of:

- Forget Gate: Decides what information to discard
- Input Gate: Decides what new information to store
- Output Gate: Produces the final hidden state

However, LSTMs require sequential processing, which makes them incompatible with OAK-D unless unrolled into a fully connected network.

Unrolling LSTM for OAK-D. To run LSTMs on OAK-D, we replace recurrence with fully connected layers, flattening the time-series input to mimic sequential learning.

$$h_T = W_o \bullet h_{t-1} + b \tag{19}$$

This removes the sequential recurrence, allowing parallel execution.

Transformer Model. Transformers use self-attention to model long-range dependencies without recurrence, allowing parallelization.

For an input sequence the Self-Attention Mechanism is computed as:

$$Q = XW_Q \tag{20}$$

$$K = XW_K \tag{21}$$

$$V = XW_V \tag{22}$$

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{23}$$

where Q, K, V are query, key and value matrices derived from input and d_k is the scaling factor. The output of a Transformer block is:

$$H = LayerNorm(FFN(SelfAttention(X)) + X) \tag{24}$$

where *FFN* is a feedforward network.

CNN Model (LSTM Alternative for OAK-D). CNNs apply 1D convolutional operations to sequential data, allowing parallel computation.

The standard convolution operation is defined as:

$$y_t = \sigma(W * X + b) \tag{25}$$

In this manner CNNs process entire time series simultaneously.

Hybrid TCN-Transformer Model. This model combines the TCN efficiency with Transformer's attention mechanisms. Using the TCNs we extract short-term temporal dependencies while using the transformer encoder we capture long-range dependencies using self-attention.

We use the TCN model to handle sequential data without recurrence, using dilated convolutions for efficiency and Transformers for handling global dependencies, allowing the model to focus on important temporal relationships.

The first part of the model applies TCN layers to extract sequential patterns. Each CNN block is computed as:

$$y_t = \sum_{i=0}^{k-1} W_i \bullet x_{t-d \bullet i} \tag{26}$$

where k is the kernel size, d is the dilation rate and W_i is the convolutional weight matrix. CNN uses residual connections:

$$H_{TCN} = ReLU(X + TCN(X)) \tag{27}$$

Which improves gradient flow and stabilizes learning.

The Transformer Encoder takes the TCN output and applies self-attention to model long-term dependencies. Each Transformer layer follows:

$$H = LayerNorm(FFN(SelfAttention(H_{TCN})) + H_{TCN}) \tag{28}$$

where *SelfAttention* is the same as in (23) and *FeedForward Network* (*FFN*) is as follows:

$$FFN(X) = ReLU(XW_1 + b_1)W_2 + b_2 \tag{29}$$

After TCN and Transformer processing, we apply global average pooling to reduce dimensionality and Dense layers or classification. The final predictions are obtained using:

$$\hat{y} = softmax(WH + b) \tag{30}$$

Hyperparameter for OAK-D Comparison. In this section we provide the Table 1 with the hyperparameter comparison of the afore-mentioned models.

3 Results and Discussions

In this section we present the results for the dimensionality reduction methods (PCA, kPCA and UMAP), training results for intention classifiers along with their execution times on OAK-D platform.

Table 1. Trained models hyperparameter comparison.

Model	Hidden units / Filters	Kernel size	Attention heads	Feedforward Dim	Dropout
LSTM	64	N/A	N/A	N/A	0.3
Transformer	N/A	N/A	4	128	0.2
CNN	32, 64	3	N/A	N/A	0.1
Hybrid TCN + Transformer	32	3	2	64	0.1

3.1 Analysis of Dimensionality Reduction Techniques

To assess the effectiveness of different dimensionality reduction techniques for skeleton-based action classification, we evaluated the classification performance of a Random Forest Classifier trained on features selected via Principal Component Analysis (PCA), Kernel PCA (kPCA), and Uniform Manifold Approximation and Projection (UMAP). The classification accuracy serves as an indicator of how well each method retains the most discriminative skeletal joints for distinguishing between actions. However, accuracy alone does not fully capture the model′s effectiveness, as it does not account for class imbalances. Therefore, we further analyze precision, recall, and F1-score, which provide deeper insight into model performance.

Retained Joints and Quantitative Results. Each dimensionality reduction method selects 10 key skeletal joints, which are subsequently used as features for classification. Table 2 presents the top joint respective classification performance metrics.

The joints retained by each method include:

- **PCA**: RightForeArm, LeftForeArm, LeftHand, RightHand, LeftArm, RightArm, LeftUpLeg, Hips, RightUpLeg, Head
- **kPCA**: HeadRotation, LeftHand, LeftForeArm, RightHand, RightForeArm, RightUpLeg, RightArm, Head, LeftArm, Neck
- **UMAP**: Hips, RightUpLeg, LeftUpLeg, Spine, RightArm, LeftArm, Head, Neck, RightForeArm, LeftForeArm

Table 2. Classification performance for each dimensionality reduction method.

Method	Accuracy [%]	Precision	Recall	F1-score
PCA	47.32	0.4664	0.4732	0.3849
kPCA	50.31	0.3967	0.5031	0.4207
UMAP	46.65	0.3797	0.4665	0.3648

PCA Performance. PCA is a widely used linear dimensionality reduction technique that projects data onto a lower-dimensional subspace while preserving maximum variance. The obtained accuracy of 47.32% suggests that the top-ranked joints selected by PCA provide some discriminative power but are insufficient for robust classification. The precision of 46.64% and recall of 47.32% indicate that PCA performs moderately well at correctly identifying actions. However, the F1-score of 38.49% suggests that the model struggles with class balance, likely misclassifying certain actions due to variance-based feature selection.

kPCA Performance. Kernel PCA extends PCA by employing non-linear kernel transformations, allowing it to map the data into a higher-dimensional feature space where complex relationships can be better captured. The highest accuracy of 50.31% suggests that non-linear transformations improve joint selection. However, despite the accuracy gain, precision drops to 39.67%, which means kPCA produces more false positives compared to PCA. The recall of 50.31% indicates that the classifier is correctly capturing a larger proportion of actual instances, but at the cost of precision. The F1-score of 42.07% suggests that, while kPCA offers better recall than PCA, it does not significantly improve class-wise balance.

UMAP Performance. UMAP is a non-linear manifold learning technique designed for preserving both local and global structures in data. Surprisingly, it resulted in the lowest accuracy of 46.65%, slightly underperforming PCA. Unlike PCA and kPCA, UMAP retained more core skeletal joints (e.g., Hips, Spine, and Neck) rather than extremities (e.g., hands). This suggests that UMAP prioritizes global pose structure over localized motion, which may not be ideal for distinguishing actions where hand gestures are crucial. The precision (37.97%) and F1-score (36.48%) are the lowest among all methods, indicating that UMAP introduces more classification noise, potentially due to its stochastic nature and sensitivity to hyperparameters.

3.2 Classification Models Results

The performance of four different neural network models—Hybrid model, Temporal Convolutional Neural Network (tCNN), Long Short-Term Memory (LSTM), and Transformer—was evaluated for intention recognition using classification reports. The comparison of these models is based on key performance metrics, namely precision, recall, and F1-score. For training purposes a train/validation split is used as proposed in [25] where the authors provide lists of sequences for training and validation. After data preprocessing training data includes ~ 1.9M samples while validation data includes ~ 600k samples. Additionally training data is split into training and test data with a ratio of 80/20 for training and testing.

Overall Performance. As presented in Table 3 the Hybrid model demonstrated the highest overall performance with an accuracy of 54.54%, outperforming the tCNN (54.54%), LSTM (48.93%), and Transformer (48.20%) models. The macro and weighted averages of the F1-score indicate that the Hybrid model provides a more balanced classification across multiple intention classes. The tCNN model exhibited similar results to the

Hybrid model, suggesting that convolutional feature extraction contributes significantly to distinguishing action categories.

The LSTM and Transformer models performed worse in comparison, with macro F1-scores of 0.2440 and 0.2584, respectively. Their lower recall values indicate that these architectures struggle to capture temporal dependencies effectively within the dataset. Specifically, Transformer-based intention recognition did not generalize well across all action categories, as indicated by the drastic drop in recall for certain actions (e.g., "Put down measuring rod" and "Take subsystem").

It is also important to note that the poor classification results can be attributed in part to the inference process during validation, which was conducted on an OAK-D device. The data was compressed to the F16 value type, leading to potential information loss that may have negatively impacted model predictions.

Table 3. Overall performance.

Model	Accuracy	Macro Avg Precision	Macro Avg Recall	Macro Avg F1-score	Weighted Avg Precision	Weighted Avg Recall	Weighted Avg F1-score
Hybrid	0.5454	0.3806	0.3517	0.3564	0.5382	0.5454	0.5383
tCNN	0.5454	0.3806	0.3517	0.3564	0.5382	0.5454	0.5383
LSTM	0.4893	0.2593	0.2484	0.2440	0.4690	0.4893	0.4740
Transformer	0.4820	0.3295	0.2555	0.2584	0.4829	0.4820	0.4734

Among individual intention categories, "Picking in front" and "Take measuring rod" achieved relatively high F1-scores across all models, with values reaching up to 0.6407 and 0.6058, respectively, in the Hybrid model. This suggests that these actions are well-represented within the dataset and are distinguishable using learned feature representations.

However, certain classes, such as "Take subsystem" and "Put down measuring rod," suffered from low recall scores across all models, indicating that these actions were either underrepresented in the dataset or lacked distinguishable patterns. The Transformer model particularly struggled with these actions, with recall values dropping as low as 0.0325.

Hybrid Model on OAK-D Performance. Along with the standard classification evaluation methods we provide the MIN, MEAN and MAX inference durations with the exception of LSTM model which lacks support on OAK-D. The measurements are acquired by running validation on OAK-D platform. The results are presented in Table 4.

Challenges in Intention Classification. The classification reports highlight the difficulty of recognizing the "No action" category, where all models reported zero precision and recall. This suggests that either the dataset lacks enough labeled examples for this class or the models fail to distinguish the absence of action from subtle movement noise.

Moreover, the LSTM model exhibited lower performance across most categories, particularly for tasks involving tool manipulation such as "Take screwdriver" and "Put

Table 4. Model-Wise Inference duration.

Model	Min duration [ms]	Mean duration [ms]	Max duration [ms]
Hybrid	5.839347839355469	7.1648098707199095	84.00130271911621
tCNN	2.999544143676758	4.713894081115723	53.99966239929199
Transformer	2.999544143676758	4.605532503128051	75.99949836730957

down screwdriver." This indicates that sequence-based learning alone is insufficient for precise intention recognition, likely due to the high variability in execution time and motion patterns for these tasks.

Model Selection Considerations. From these results, it is evident that a hybrid approach incorporating both convolutional and recurrent layers provides the most balanced performance across diverse action categories. The Hybrid and tCNN models outperform the LSTM and Transformer models, suggesting that localized spatial-temporal feature extraction plays a crucial role in improving classification performance.

Future improvements may involve refining the dataset with better-balanced class distributions and employing ensemble learning strategies to enhance classification robustness across underrepresented action categories.

4 Conclusions and Future Work

This study presented an in-depth analysis of dimensionality reduction methods (PCA, kPCA, UMAP) and intention recognition models (Hybrid model, Temporal CNN, LSTM, and Transformer) within an industrial human-robot collaboration (HRC) context. The results highlight the trade-offs between feature selection techniques and classifier architectures, particularly in terms of classification performance, computational efficiency, and real-time execution feasibility.

From the dimensionality reduction evaluation, kPCA achieved the highest accuracy (50.31%), suggesting that non-linear feature extraction improves joint selection. However, the observed precision drop indicates a higher rate of false positives, which poses challenges for action classification in complex environments. In contrast, UMAP prioritized global skeletal structures (e.g., Hips, Spine, and Neck), leading to a lower classification performance (46.65%), likely due to its stochastic nature.

Among the intention recognition models, the Hybrid model and tCNN outperformed LSTM and Transformer architectures, achieving an accuracy of 54.54%. The Transformer model (48.20%) showed inconsistent recall, indicating difficulties in generalizing across action categories. The LSTM model (48.93%) struggled with temporal dependencies, particularly for tasks involving tool manipulation, suggesting that sequence-based learning alone is insufficient for robust action recognition in industrial settings. Furthermore, inference testing on the OAK-D platform highlighted computational challenges, with LSTM lacking hardware support and Transformer models experiencing significant recall degradation in complex actions.

While this study provides valuable insights into feature selection and intention classification, several challenges remain that require further investigation. Future research will focus on improving Transformer-based models by addressing their limitations in intention recognition through domain-adaptive fine-tuning and architectural enhancements. Integrating hybrid attention mechanisms that combine self-attention with CNN-based spatial encodings may enhance the model′s ability to generalize across diverse action categories.

Another critical aspect is improving real-time processing, as inference latency remains a challenge for industrial applications. Future work will explore quantization-aware training and optimized model compression techniques to ensure efficient deployment on edge devices like the OAK-D platform. Additionally, the reliance on skeletal joint tracking limits contextual awareness in current models. To overcome this, future research will investigate the integration of RGB-D data, LiDAR-based pose estimation, and multimodal sensor fusion, which could significantly enhance human intention recognition in highly dynamic industrial environments.

Expanding the dataset and addressing class imbalance will also be a priority, as the low recall for certain action categories indicates an uneven distribution of training data. Future research will employ data augmentation techniques, active learning strategies, and synthetic data generation to improve recognition rates for underrepresented actions. Finally, developing adaptive and continual learning frameworks will be essential for handling the variability in industrial human actions. Implementing these strategies will allow systems to incrementally improve their performance over time while remaining flexible to new task scenarios, ensuring that human-robot collaboration remains efficient and scalable in real-world applications.

By addressing these challenges, next-generation action recognition models will be more scalable, computationally efficient, and robust, ensuring seamless human-robot collaboration in complex industrial environments.

Acknowledgments. This research is partially supported by Visage Technologies AB and by the Croatian Science Foundation under the project "Affective Multimodal Interaction based on Constructed Robot Cognition—AMICORC (UIP-2020-02-7184)."

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Aggarwal, J.K., Xia, L.: Human activity recognition from 3D data: a review. Pattern Recogn. Lett. **48**, 70–80 (2014). https://doi.org/10.1016/j.patrec.2014.04.011
2. Cao, Z., Hidalgo, G., Simon, T., Wei, S.E., Sheikh, Y.: OpenPose: realtime multi-person 2D pose estimation using part affinity fields. IEEE Trans. Pattern Anal. Mach. Intell. **43**(1), 172–186 (2019). https://doi.org/10.1109/TPAMI.2019.2929257
3. Wang, J., Chen, Y., Hao, S., Peng, X., Hu, L.: Deep learning for sensor-based activity recognition: a survey. Pattern Recogn. Lett. **119**, 3–11 (2019). https://doi.org/10.1016/j.patrec.2018.02.010

4. Pham, H.H., Khoudour, L., Crouzil, A., Zegers, P., y Velastin, S.A.: Exploiting deep residual networks for human action recognition from skeletal data. Comput. Vis. Image Underst. **170**, 51–66 (2018). https://doi.org/10.1016/j.cviu.2018.03.003
5. Nguyen, H.-C., Nguyen, T.-H., Scherer, R., Le, V.-H.: Deep learning for human activity recognition on 3d human skeleton: survey and comparative study. Sensors **23**(11), 5121 (2023). https://doi.org/10.3390/s23115121
6. Clarke, R., et al.: The properties of high-dimensional data spaces: implications for exploring gene and protein expression data. Nat. Rev. Cancer **8**(1), 37–49 (2008). https://doi.org/10.1038/nrc2294
7. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. J. Mach. Learn. Res. **3**, 1157–1182 (2003). https://doi.org/10.1162/153244303322753616
8. Arulananth, T.S., Balaji, L., Baskar, M., Anbarasu, V.: PCA-based dimensional data reduction and segmentation for DICOM images. Neural. Process. Lett. (2023). https://doi.org/10.1007/s11063-020-10391-9
9. Upreti, D., Kim, H., Yang, E.: Defending against label-flipping attacks in federated learning systems with UMAP. Research Square (2022). https://doi.org/10.21203/rs.3.rs-1984301/v1
10. Rezazadeh, N., Annaz, F., Jabbar, W.A.: A transfer learning approach for mitigating temperature effects on wind turbine blades damage diagnosis. Struct. Health Monit. (2025). https://doi.org/10.1177/14759217241313350
11. Alnafisah, K.H., Ranjan, A., Sahu, S.P.: Machine learning for automated classification of lung collagen in a urethane-induced lung injury mouse model. Biomed. Opt. Express (2024). https://doi.org/10.1364/BOE.474361
12. Bairagi, A., Das, B.P., Dey, A.: A nonlinear projection-based human emotion recognition approach employing face point data. IEEE Int. Conf. Syst. Man Cybern. (2024). https://doi.org/10.1109/SMC.2024.10468402
13. Khan, A.R., Ahamad, S.S., Mishra, S., Khan, M.A.R.: FinSafeNet: securing digital transactions using optimized deep learning and multi-kernel PCA (MKPCA) with Nyström approximation. Sci. Rep. (2024). https://doi.org/10.1038/s41598-024-76214-2
14. Samerei, S.A., Aghabayk, K., Montella, A.: Analyzing pile-up crash severity: insights from real-time traffic and environmental factors using ensemble machine learning and shapley additive explanations method. Safety **10**(1), 22 (2024). https://doi.org/10.3390/safety10010022
15. Guo, Z.J., et al.: Interpretable machine learning models based on shapley additive explanations for predicting the Risk of cerebrospinal fluid leakage in lumbar fusion surgery. Spine (2024). https://doi.org/10.1097/BRS.0000000000005087
16. Orsag, L., Stipančić, T., Koren, L.: Towards a safe human-robot collaboration using information on human worker activity. Sensors (2022). https://doi.org/10.3390/s23031283
17. Orsag, L., Stipančić, T., Koren, L., Posavec, K.: Human intention recognition for safe robot action planning using head pose. HCI Conf. Proc. (2022). https://doi.org/10.1007/978-3-031-17618-0_23
18. Wang, X., Chen, Y., Wang, F., Zheng, K., Zong, Y.: Mpct-LSTM: a lightweight human activity recognition model for 3D point clouds in millimeter-wave radar. SSRN Pap. (2024). https://doi.org/10.2139/ssrn.5081681
19. Zhang, Q., et al.: Leader–follower: Human-centered intention-guided controller for novel SuperLimb with application to load-carrying scenarios. Sci. Chin. (2025). https://doi.org/10.1007/s11431-024-2806-8
20. Chen, D., Chen, M., Wu, P., et al.: Two-stream spatio-temporal GCN-transformer networks for skeleton-based action recognition. Sci. Rep. **15**, 4982 (2025). https://doi.org/10.1038/s41598-025-87752-8
21. Vaswani, A.: Attention is all you need. Adv. Neural Inf. Process. Syst. **30** (2017). https://doi.org/10.48550/arXiv.1706.03762

22. Mazzia, V., Angarano, S., Salvetti, F., Angelini, F., Chiaberge, M.: Action transformer: a self-attention model for short-time human action recognition. Pattern Recogn. **124** (2021). https://doi.org/10.48550/arXiv.2107.00606
23. Guan, S., et al.: Classification of cigarette types using computer vision: an analysis of smoke aggregation features. IEEE Access **13**, 11836–11845 (2025)
24. Islam, M.M.: AI-Driven multi-purpose platform for smart healthcare systems. Waterloo Research Space (2025). https://uwspace.uwaterloo.ca/items/9c329461-b19c-4267-8841-b4e5e050d290
25. Dallel, M., Havard, V., Baudry, D., Savatier, X.: InHARD industrial human action recognition dataset in the context of industrial collaborative robotics. Zenodo **2020–09**, 30 (2020). https://doi.org/10.5281/zenodo.4003541

First Steps Toward the Agile Integration of Information Architecture into a User-Centered Development Process

Luis Rojas[1(✉)], Katrina Sorbello[2], Vicente Marin[1], Fei Luo[3], John W. Castro[4], Avon Collis[5], and Pablo Schwarzenberg[6]

[1] Facultad de Ingeniería, Universidad San Sebastián, Bellavista 7, 8420524 Santiago, Chile
luis.rojasp@uss.cl, vmarina3@correo.uss.cl
[2] The Stella Way, 17 Enford Street, Hillcrest, QLD, Australia
kat@thestellaway.com
[3] School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China
luof@ecust.edu.cn
[4] Departamento de Ingeniería Informática y Ciencias de La Computación, Universidad de Atacama, Chile y Centro Interuniversitario de Envejecimiento Saludable (CIES), Copiapó, Chile
john.castro@uda.cl
[5] Relevate, 152 Woogaroo Street, Forest Lake, QLD, Australia
avon.collis@relevate.au
[6] Factulad de Ingeniería, Universidad Andrés Bello, Santiago, Chile
pablo.schwarzenberg@unab.cl

Abstract. The Information Architecture (IA) is a term of relevance today that makes sense in the development of applications and information systems. In fact, a crucial aspect in IA is that it should promote usability while being flexible enough to match the dynamic information needs of the business environments. Currently, there are various methodological proposals to the IA. However, most of them present difficulties to prioritize requirements on content usability, and they do not fully cover the development process, focusing only on some of their phases. Likewise, most of the proposals require a significant effort to adapt to complex environments. This paper addresses the problem of adapting the activities of the IA for the development of interactive applications improving the satisfaction of end-users and promoting an agile integration of the IA into a user-centered development framework. Specifically, in this work in progress we propose the initial guidelines of an agile methodology, called UXAgileIA, to develop the Information Architecture through a user-centered development process, providing different activities.

Keywords: Development Process · Information Architecture · User-Centered Design · Agile · Usability

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 87–96, 2025.
https://doi.org/10.1007/978-3-031-93539-8_6

1 Introduction

The term Information Architecture (IA) is defined as the art and discipline for the design, analysis, organization, arrangement and structuring of information [1, 2]. This highlights the need for good management and provision of information, being also accessible by end-users and in-line with the organization's objectives that provide this information.

In this context, the symbiosis between IA and User Centered Design (UCD) is essential to be able to propose an analysis and design of content in accordance with user requirements and usability. [3]. In fact, there are different works [4, 5] that emphasize the importance of focusing IA activities on the needs of end-users. However, and despite its imminent importance, these works do not identify or describe the necessary stages in the IA deployment process, making it difficult to establish the requirements for an analysis and design of usable content from the early phases of the project. Likewise, organizations require a correct IA to efficiently adapt to the agile environments in which they operate [6], which leads to the need to have new methodological proposals for its development [7].

Currently, there are several methodological proposals for IA [1, 8–11]. However, these are based on traditional models of the development process and require considerable effort to adapt. Therefore, it is difficult to carry them out in changing environments, as an agile and rapid response is often required.

In general, agile methodologies arise in response to the need to quickly adapt to changing environments. However, these methodologies also present difficulties when incorporating UCD in the agile development process [12–15], this being a primary factor to efficiently integrate IA. Several studies have addressed the need to involve end-users in agile development [16–21], providing specific recommendations, but without providing a comprehensive or complete vision. Various authors [16, 22] have reported these drawbacks as an incentive to create new agile methodologies or approaches to integrate IA activities within a user-centered development process.

This paper is organized as follows: Sect. 1 introduced the motivation for our work. Section 2 reviews related work. Section 3 describes our proposed agile methodology. Section 4 presents a case study. Finally, Sect. 5 presents conclusions and future work.

2 Related Work

A series of proposals on IA have been compiled [1, 8–11, 23–27], according to their popularity and broad usage. These proposals are sufficiently representative to address the most current and relevant IA methodologies. On the other hand, a criteria set has been defined to compare the IA proposals. These criteria allow having a common framework to analyze the characteristics of these proposals and verify if they allow the integration of IA directed by user expectations, considering the changes in the requirements. The criteria to be used are described below:

The *User-Centered* criterion is used to identify whether the proposals design their products focused on the needs of end-users. This criterion can take *Yes, Partially* or *No* values to indicate that the proposals are fully, partially or not user-centered, respectively. The *IA Elements Covered* criterion is used to identify the aspects of the IA that

proposals attempt to cover. According to Erlin et al. [28] the aspects of the IA that are involved in the development of a product correspond to: navigation (N), organization (O), labeling (L), search and recovery (SR), and information presentation (IP). The *Level of Description* criterion is used to identify the level of detail in which methodologies for the IA development are described. This criterion can take *High, Average* or *Low* values to indicate that the proposals are described with sufficient, moderate or general detail, respectively. The *Scope of the Proposal* criterion is used to identify the lifecycle phases of IA development that proposals attempt to cover. The considered phases are analysis (A), design (D), evaluation (E), implementation (I), and management (M). The phases that are incompletely covered (without providing or prescribing the necessary information to carry them out) are pointed out with an (-). Finally, the *Flexibility and Adaptability* criterion is used to identify whether methodologies are able to adapt to changing environments and respond quickly. In order to assess this criterion, IA methodologies are analyzed to know whether they require extensive planning or have a development process model (i.e., waterfall) that needs numerous controls and policies/standards to be implemented.

Table 1. Summary of the comparative analysis of the selected study sources.

Method	User-Centered	IA Elements Covered	Level of Descr	Scope of the Prop.	Flex. And Adpt.
[9]	No	O	Low	A (-), D (-)	No
[29]	Partially	N, O, L, SR, IP	High	A, D, I (-), M (-)	No
[8]	Partially	N, O, L	Low	A, D	No
[11]	Partially	O, IP,	Low	A (-), D (-)	No
[24]	No	O, IP	Low	A, D (-), E (-)	No
[25]	Partially	N, O, IP	Low	A, D	No
[26]	Partially	N, O, L	Low	A, D	No
[10]	Yes	N, O, L, SR, IP	High	A, D, E, I, M	No
[23]	Partially	N, O, L	Low	A, D, E (-)	No
[27]	Partially	N, L	Average	E	No

As can be seen in Table 1, most of the proposals do not cover all phases of the IA life cycle, focusing mainly on the analysis and design phases. In fact, only proposals [27] and [10] present recommendations for activities and techniques for the evaluation of IA (Criterion Scope of the Proposal: Evaluation). Similarly, proposals [24] and [23] also suggest carrying out their evaluation. However, these proposals are described superficially (-) , without providing enough information to carry out such an evaluation. Likewise, the methodology [10] is the only one that provides a user-centered approach to integrate IA (User-Centered Criterion: Yes). On the contrary, in most of the proposals this aspect

is carried out partially, [1, 8, 11, 23, 25–27, 29], that is, the end-users are considered in the initial phases and not throughout the model development process. Likewise, the methodologies [1] and [10] are the only ones that have an adequate level of description of the proposals (Description Level Criteria: High) and cover all the elements of the IA (Criteria Elements Covered of IA) in the different phases of its life cycle (Criterion Scope of the Proposal). Regarding the Flexibility and Adaptability criterion, the methodology [25] is the only one that suggests integrating IA in an agile way. However, this proposal does not identify the necessary details (process model, activities, tasks, relationships, among others) to carry out a correct integration of IA in changing environments.

3 A Preliminary Agile Proposal to Integrate Information Architecture into a User-Centered Development Process

An agile methodology, called UXAgileIA, is proposed to integrate IA through a user-centric development process (see Fig. 1). This methodology is based on the Scrum framework and includes roles, events, artifacts, and associated rules. This approach is a combination of the recommendations and the specific practices to integrate the UCD in the Scrum process. The most common practices correspond to:

- Generate a definition of "Done" for the tasks related to the UCD [30, 31];
- Incorporate elements in the Product List that include a greater awareness of usability [32, 33];
- Utilization of informal ways to involve end-users in the evaluation [30, 34];
- Carry out a contextual investigation, before starting the software development [30, 35];
- Establish close collaboration between UCD specialists and developers in Scrum teams [32–34];
- Assign explicit responsibility from the end-user perspective [22];
- Systematize the process of participation of end-users [22].

As shown in Fig. 1, the team is made up following the same traditional Scrum structure (a Product Owner, the Development Team and a UXAgileIA Master), but the Information Architect is incorporated with a primary role, in order to direct the Contextual Research, support the management of the Product Backlog, ensure the UCD and encourage the participation of end-users.

UXAgileIA is based on the establishment of a Contextual Research as the starting point of the project, with the aim of studying and analyzing the needs of end-users (Activities 1.1 and 1.2), as well as prioritizing content (Activity 1.3), a problem reported by [22]. In this way, knowledge is acquired and the bases are provided to establish an overview of the project (Activity 1.4), as recommended in [35] and successfully implemented in [30].

Subsequently, in the Requirements Management, the requirements that are included in the Product List are prioritized collaboratively (Activities 2.1, 2.2, 2.3 and 2.4), according to the priorities obtained in the Contextual Research. After the Requirements Management, the Sprint Planning is carried out, through the Sprint Planning Meeting event. In this event, the work to be done in each project Sprint is defined (Activities 3.1 and

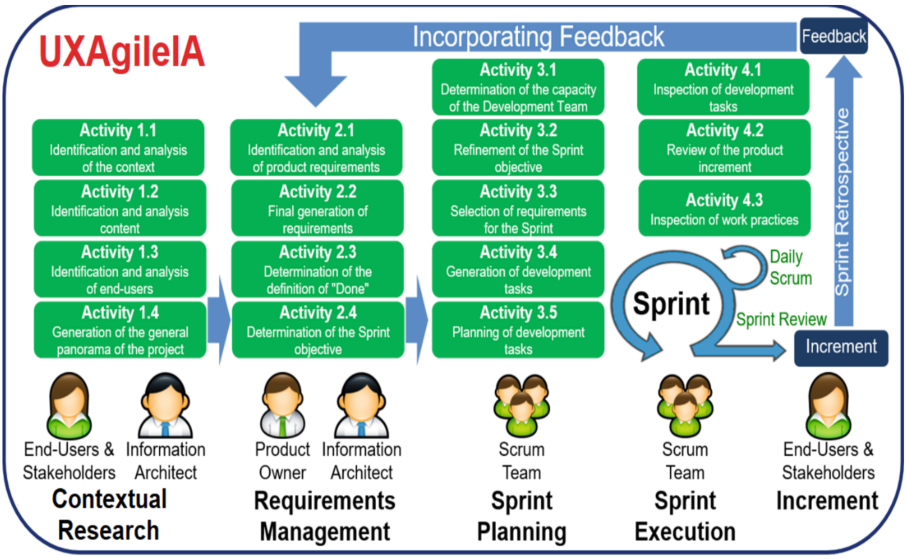


Fig. 1. Activities of the UXAgileIA methodology.

3.2). Therefore, the highest priority requirements from the Product Backlog are selected for the Sprint Backlog (Activity 3.3). Such a list is made up of the requirements, selected from the Product Backlog, together with a plan to carry them out.

Following that, the Development Team divides these requirements into specific tasks (Activity 3.4), which are developed during the Sprint Execution (Activity 3.5). Then, during the Sprint Execution, the Daily Scrum meetings are held (Activity 4.1). These meetings are held to review and adjust the daily work, done by the Development Team, and also to manage individual development tasks, in order to maintain visibility and consistency with the big picture.

The Sprint Review meeting is held at the end of the Sprint Execution, with the participation of end-users and other secondary stakeholders (Activity 4.2), which could also be beneficial according to [22]. The objective of this meeting is to analyze the potential increase of the product, generated by the Development Team. Finally, the Sprint Retrospective meeting is held, so that the UXAgileIA team can analyze the work practices used during the iteration, and look for improvement opportunities (Activity 4.3). Likewise, it is suggested to reflect on the results of these ceremonies immediately after the evaluations [30] and apply different methods to successfully carry out each of the user-centered evaluations [34].

3.1 Description of the UXAgileIA Integration Framework

In this section, we detail the activities of the UXAgileIA methodology, as illustrated in Fig. 1. As shown, the activities are organized into four main groups: Contextual Investigation, Requirements Management, Sprint Planning, and Inspection and Continuous Improvement.

The activity groups of Contextual Investigation, Requirements Management, and Sprint Planning correspond directly to the distinct phases proposed in the UXAgileIA methodology. On the other hand, the Inspection and Continuous Improvement activity group encompasses activities related to the various evaluations carried out in UXAgileIA to promote User-Centered Design (UCD) and the involvement of end users. The structure of the UXAgileIA activities is as follows:

Contextual Investigation Activity Group. The purpose of this activity group is to gain insights into the priorities of end users, business value, and content within the project's context. It also establishes the foundation for creating a general overview of the project so that the end users' perspective can be discussed, described, and integrated into the UXAgileIA process model. The activities within this group are:

- *Activity 1.1: Context Identification and Analysis.* This activity analyzes the internal and external characteristics of the organization, focusing on economic, technological, legal, and socio-cultural aspects.
- *Activity 1.2: Content Identification and Analysis.* In this activity, the contents related to the system and the project′s content-related priorities are identified and analyzed.
- *Activity 1.3: End-User Identification and Analysis.* This activity identifies and analyzes all end users of the system to determine their main characteristics and needs.
- *Activity 1.4: General Project Overview Generation.* This activity generates an integrated and global vision of the product, taking into account content priorities (from Activity 1.2), business value (from Activity 1.1), and end-user priorities (from Activity 1.3).

Requirements Management Activity Group. The goal of this activity group is to identify, analyze, prioritize, and estimate project requirements. It builds upon the knowledge obtained in the Contextual Investigation Activity Group to ensure that project requirements reflect business value, end-user priorities, and content priorities. The activities in this group are:

- *Activity 2.1: Project Requirements Identification and Analysis.* This activity identifies and analyzes the functional requirements of the system by considering the end users and business value defined in Activity 1.4.
- *Activity 2.2: Final Requirements Generation.* Here, the requirements identified in Activity 2.1 are described, analyzed, estimated, and prioritized to create the Product Backlog.
- *Activity 2.3: Definition of "Done" Determination.* This activity establishes the definition of "Done" for the requirements identified in Activity 2.2 to prepare for the next Sprint.
- *Activity 2.4: Sprint Goal Determination.* An initial Sprint goal is identified and defined in this activity to start the Sprint Planning phase with a clear focus on end-user priorities.

Sprint Planning Activity Group. This group aims to identify, estimate, and plan development tasks required to implement the selected requirements for the Sprint (identified in the Requirements Management Activity Group). The activities in this group are:

- *Activity 3.1: Development Team Capacity Determination.* This activity assesses the velocity and resource availability of the Development Team to establish the Sprint duration.
- *Activity 3.2: Sprint Goal Refinement.* The initial Sprint goal identified in Activity 2.4 is finalized for use in guiding Sprint Execution.
- *Activity 3.3: Requirement Selection for the Sprint.* This activity involves selecting requirements from Activity 2.2 for the Development Team to execute, forming part of the Sprint Backlog.
- *Activity 3.4: Development Task Generation.* Specific development tasks required to implement the selected requirements are identified and defined in this activity.
- *Activity 3.5: Development Task Planning.* The execution of development tasks identified in Activity 3.4 is organized to ensure incremental development with varying levels of fidelity in AI deliverables.

Inspection and Continuous Improvement Activity Group. This group focuses on examining and evaluating the development tasks and work practices used during the Sprint to foster UCD and end-user involvement. The activities in this group are:

- *Activity 4.1: Development Task Inspection.* This activity reviews the status of development tasks (identified in Activity 3.4 and planned in Activity 3.5) and addresses issues faced by the Development Team, ensuring compliance with end-user acceptance criteria.
- *Activity 4.2: Product Increment Review.* The product increment generated by the Development Team during the Sprint Execution is reviewed and inspected with end-user involvement.
- *Activity 4.3: Work Practices Inspection.* This activity reflects on the completed Sprint to identify improvements for the next Sprint and opportunities related to end-user involvement.

4 Conclusions and Future Work

A work in progress has been proposed for an agile integration of Information Architecture (IA) into a user-centric development process that allows to efficiently align the needs of end-users and the business with the interactive applications development. The main contribution of this paper consists of a methodology, called UXAgileIA, which allows the integration of IA through an agile user-centered development process. In this methodology, the end-user's perspective is analyzed, described and considered during the Product Backlog management process, driven by contextual research. Also, Requirements implementation is driven by usability, IA, and business value priorities through Sprint planning, driven by IA deliverables. Finally, Compliance with end-user acceptance criteria is checked early in the Sprint, and product increments are evaluated through end-user-driven inspection and continuous improvement processes.

As future work, we intend to refine the methodology by specifying detailed guidelines for input and output products, techniques, and tasks. These enhancements will ensure a more precise and successful implementation of the proposal in practice. Additionally, we

expect to create a Web tool, featuring a project management system driven by our proposal. The tool will include different functionalities to support the agile and user-centered development of the IA, including usability and user-centric concerns. Furthermore, we plan to implement the proposed methodology across diverse software projects to evaluate its effectiveness and efficiency. This will include assessing the satisfaction levels of both professional teams and end users. Likewise, we want to formally evaluate user experience and usefulness requirements [36, 37] of the proposal with experts in IA.

References

1. Rosenfeld, L., Morville, P., Arango, J.: Information architecture: for the web and beyond. O'Reilly Media (2015)
2. Guizani, M.: A decade of information architecture in HCI: a systematic literature review. ArXiv Prepr. ArXiv220213412 (2022)
3. Abelein, U., Paech, B.: Understanding the influence of user participation and involvement on system success a systematic mapping study. Empir. Softw. Eng. **20**(1), 28–81 (2015). https://doi.org/10.1007/s10664-013-9278-4
4. de Lima Salgado, A., Pereira, F.H., Freire, A.P.: User-centred design and evaluation of information architecture for information systems. In: Handbook of Research on Information Architecture and Management in Modern Organizations IGI Global, pp. 219–236 (2015)
5. Alshaheen, R., Tang, R.: User experience and information architecture of selected national library websites: a comparative content inventory, heuristic evaluation, and usability investigation. J. Web Librariansh. **16**(1), 31–67 (2022)
6. Morris, A., Setchi, R., Prickett, P.: Product change management and future information architectures. In: Setchi, R., Howlett, R., Liu, Y., Theobald, P. (eds.) Sustainable Design and Manufacturing 2016. SDM 2016. Smart Innovation, Systems and Technologies, vol. 52, pp. 241−250. Springer, Cham. (2016). https://doi.org/10.1007/978-3-319-32098-4_21
7. Martin, A., Dmitriev, D., Akeroyd, J.: A resurgence of interest in information architecture. Int. J. Inf. Manag. **30**(1), 6–12 (2010)
8. Huang, X., Zhang, Y.: Designing a teaching website for early warning technology support specialty based on information architecture. J. Contemp. Educ. Res. **6**(1) (2022)
9. do Rocio Maciel, D., Cerutti, D.M.: Usability, information architecture and semiotic engineering applied to the writing of support systems. Iberoam. J. Appl. Comput. **5**(2) (2016)
10. Reichenauer, A.: LUCIA: development of a comprehensive information architecture process model for websites. Doctoral dissertation, University of Regensburg, Germany (2005)
11. Zhang, X., Zhou, X., Liu, J., Xu, J.: 10-Step information architecture on wanfang med online. In: Future Information Technology-II, pp. 1–9. Springer (2015)
12. Choudhary, B., Rakesh, S K.: An approach using agile method for software development. In: Innovation and Challenges in Cyber Security (ICICCS-INBUSH), 2016 International Conference on, pp. 155–158. IEEE (2016)
13. Elbanna, A., Sarker, S.: The risks of agile software development: learning from adopters. IEEE Softw. **33**(5), 72–79 (2016)
14. Salah, D., Paige, R.F., Cairns, P.: A systematic literature review for agile development processes and user centred design integration. In: Proceedings of the 18th international conference on evaluation and assessment in software engineering, p. 5. ACM (2014)
15. Zorzetti, M., Signoretti, I., Salerno, L., Marczak, S., Bastos, R.: Improving agile software development using user-centered design and lean startup. Inf. Softw. Technol. **141**, 106718 (2022)

16. Brhel, M., Meth, H., Maedche, A., Werder, K.: Exploring principles of user-centered agile software development: a literature review. Inf. Softw. Technol. **61**, 163–181 (2015)
17. Isomursu, M., Sirotkin, A., Voltti, P., Halonen, M.: User experience design goes agile in lean transformation-a case study. In: 2012 Agile Conference, pp. 1−10. IEEE (2012)
18. Kuusinen, K., Mikkonen, T., Pakarinen, S.: Agile user experience development in a large software organization: good expertise but limited impact. In: Winckler, M., Forbrig, P., Bernhaupt, R. (eds.) Human-Centered Software Engineering. HCSE 2012. Lecture Notes in Computer Science, vol. 7623, pp. 94–111. Springer, Berlin, Heidelberg (2012). https://doi.org/10.1007/978-3-642-34347-6_6
19. Losada, B., Urretavizcaya, M., Fernández-Castro, I.: A guide to agile development of interactive software with a user objectives-driven methodology. Sci. Comput. Program. **78**(11), 2268–2281 (2013)
20. Memmel, T., Gundelsweiler, F., Reiterer, H.: CRUISER: a cross-discipline user interface and software engineering lifecycle. In: International Conference on Human-Computer Interaction, pp. 174–183. Springer (2007)
21. Pilz, J., Deutschländer, J., Thomaschewski, J., Schön, E.M.: Integrating agile human-centered design with lean UX and Scrum. In: WEBIST, pp. 467–473 (2021)
22. Cajander, Å., Larusdottir, M., Gulliksen, J.: Existing but not explicit-the user perspective in scrum projects in practice. In: IFIP Conference on Human-Computer Interaction, pp. 762–779. Springer (2013)
23. Lamar, L.: Introduction to a user interface design/information architecture process for Web sites. In: Professional Communication Conference, 2001. IPCC 2001. Proceedings. IEEE International, pp. 185–197. IEEE (2001)
24. Nurcahyanti, W.E.: Information architecture assessment of BPS headquarter official website. In: Information Technology Systems and Innovation (ICITSI), 2014 International Conference on, pp. 177–182. IEEE (2014)
25. Picchi, A.: Pro iOS web design and development: HTML5, CSS3, and JavaScript with Safari (2011)
26. Sharlin, M., Tu, E., Bartus, T.: Guide to creating website information architecture and content. Princet. Univ. (2009)
27. Toub, S.: Information architecture: a practical guide to assessing Web site organization (2000)
28. Yunus, Y.A., Rahman, A.A.: The evolution of information architecture. In: 2008 International Symposium on Information Technology, pp. 1–6. Kuala Lumpur, Malaysia, IEEE (2008). https://doi.org/10.1109/ITSIM.2008.4631890
29. Rosenfeld, L., Morville, P.: Information architecture for the world wide web, 3rd Edition. O'Reilly Media (2006)
30. Felker, C., Slamova, R., Davis, J.: Integrating UX with scrum in an undergraduate soft-ware development project. In: Proceedings of the 43rd ACM Technical Symposium on Computer Science Education, pp. 301–306. ACM (2012)
31. Kniberg, H.: Scrum and XP from the Trenches. Lulu. com (2015)
32. Budwig, M., Jeong, S., Kelkar, K.: When user experience met agile: a case study. In: CHI'09 Extended Abstracts on Human Factors in Computing Systems, pp. 3075–3084. ACM (2009)
33. Kuusinen, K.: Improving UX work in scrum development: a three-year follow-up study in a company. In: International Conference on Human-Centred Software Engineering, pp. 259–266. Springer (2014)
34. Lárusdóttir, M.K., Cajander, Å., Gulliksen, J.: The big picture of UX is missing in Scrum projects. In: Proceedings of the International Workshop on the Interplay between User Experience (UX) Evaluation and System Development (I-UxSED), pp. 49–54 (2012)
35. Rannikko, P.: User-centered design in agile software development (2011)

36. Rojas, L., Olivares-Rodríguez, C., Alvarez, C., Campos, P.G.: OurRank: a software requirements prioritization method based on qualitative assessment and cost-benefit prediction. IEEE Access (2022)
37. Rojas, L., Macías, J.: Toward collisions produced in requirements rankings: a qualitative approach and experimental study. J. Syst. Softw. **158**, 110417 (2019)

LLM-Driven Augmented Reality Puppeteer: Controller-Free Voice-Commanded Robot Teleoperation

Yuchong Zhang(✉), Bastian Orthmann, Michael C. Welle, Jonne Van Haastregt, and Danica Kragic

KTH Royal Institute of Technology, Stockholm, Sweden
{yuchongz,orthmann,mwelle,dani}@kth.se

Abstract. The integration of robotics and augmented reality (AR) presents transformative opportunities for advancing human-robot interaction (HRI) by improving usability, intuitiveness, and accessibility. This work introduces a controller-free, LLM-driven voice-commanded AR puppeteering system, enabling users to teleoperate a robot by manipulating its virtual counterpart in real-time. By leveraging natural language processing (NLP) and AR technologies, our system—prototyped using Meta Quest 3—eliminates the need for physical controllers, enhancing ease of use while minimizing potential safety risks associated with direct robot operation. A preliminary user demonstration successfully validated the system's functionality, demonstrating its potential for safer, more intuitive, and immersive robotic control.

Keywords: LLM-driven · Controller-free · AR puppeteer

1 Introduction

Robotics has rapidly evolved over recent years, seamlessly integrating into various aspects of daily lives [64], including service industries [14,19,24], healthcare [11,39,42], and social scenarios [3,4,46]. This growing presence highlights the need for effective and intuitive human-robot interaction (HRI) [15,18,58] and collaboration [13,20,26]. A key enabler of this evolution is the integration of robotics with augmented reality (AR), a technology that superimposes virtual elements onto the real world to create immersive, context-aware interfaces [60,61]. By bridging the gap between physical and digital realms, AR enhances usability, intuitiveness, and accessibility in human-computer interaction (HCI) [59,65]. Through its ability to present users with additional, contextually relevant information, AR provides transformative opportunities for advancing robotic teleoperation, situational awareness, and task performance [8,12,41].

Significant prior research has explored the application of AR in robotic teleoperation across diverse domains. These studies underscore AR's potential to

Y. Zhang and B. Orthmann—Contributed equally to this work.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 97–112, 2025.
https://doi.org/10.1007/978-3-031-93539-8_7

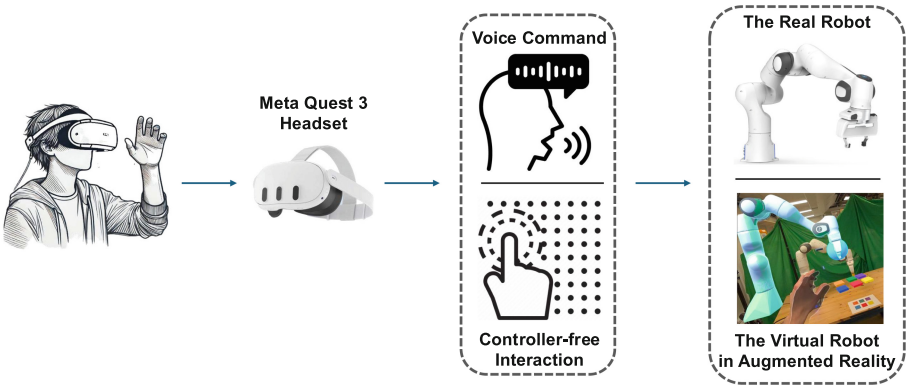


Fig. 1. An overview of our proposed LLM-driven, controller-free, voice-commanded AR robotic puppeteering system. The system operates within an AR environment, utilizing the Meta Quest 3 HMD. Users can interact seamlessly with the virtual robot, modeled identically after the real Franka robot arm using only hand gestures, while voice commands are integrated to enable intuitive control of the virtual robot, which in turn governs the real robot.

improve user experience, enhance task execution, and bolster situational awareness - key components in achieving effective HRI [2,23,37,51]. Among these contributions, the work of [48] stands out as particularly noteworthy. They introduced the first controller-based AR system for virtual robot teleoperation, enabling users to manipulate a virtual robot using physical controllers. Their innovative approach allowed the physical robot to replicate the trajectory of its virtual counterpart, laying a foundation for further advancements in AR-driven robotic teleoperation (Fig. 2).

In parallel with the evolution of AR technologies, large language models (LLMs) have emerged as a transformative force across various fields, including robotics [55,56]. Renowned for their superior natural language processing (NLP) capabilities, LLMs have been leveraged to enhance both robotic manipulation and interaction experiences [17,25]. Recently, in the realm of social robotics, LLMs facilitate intelligent conversational robots, enabling more natural and meaningful human-robot communication [28,45]. Similarly, in robotic manipulation, these models have demonstrated their potential to improve task performance by providing advanced reasoning and decision-making capabilities [9,10,29]. The integration of LLMs into robotic systems represents a promising avenue for redefining how humans interact with robots, paving the way for more intuitive, efficient, and accessible solutions.

Building on these advancements, this work introduces the LLM-driven AR Puppeteer system, a novel teleoperation framework that combines the power of LLMs and AR to enable voice-commanded, controller-free interaction with robots, as shown in Fig. 1. Expanding upon the foundation established by [48], our system eliminates the need for physical controllers, instead allowing users

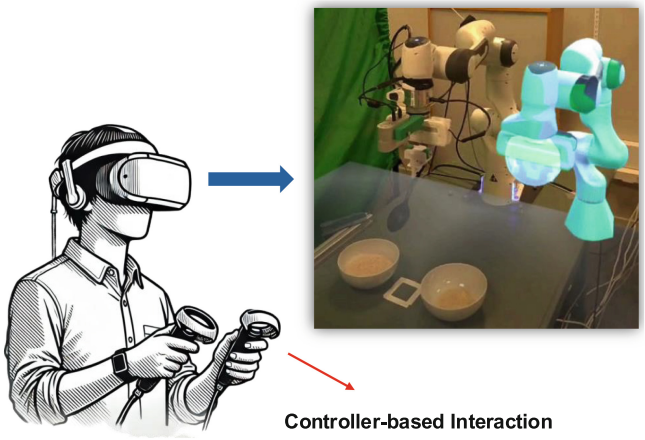


Fig. 2. The overview of the controller-based baseline AR robotic puppeteering system proposed by [48].

to intuitively manipulate a virtual robot through natural language instructions. These voice commands are processed in real-time by LLMs and seamlessly translated into actions executed by the physical robot. By replacing traditional controllers with voice-driven commands, the proposed system significantly enhances accessibility, making robotic teleoperation more inclusive for non-expert users and individuals with varying physical abilities. Our system is realized via Meta Quest 3, a new generation of optical see-through (OST) head-mounted display (HMD).

The fusion of cutting-edge NLP capabilities with immersive AR technologies offers an innovative approach to addressing challenges in HRI and HCI by introducing a controller-free, voice-driven interaction model supported by LLMs. This integration not only provides a groundbreaking framework for revolutionizing robotic teleoperation but also fosters a more natural and intuitive way for users to interact with robotic systems. By addressing barriers related to accessibility and ease of use, this human-centric approach emphasizes inclusivity and simplicity, broadening the reach of robotic applications across diverse domains [22,52,57]. For instance, in industrial settings, workers can remotely command robots to perform complex tasks without specialized training [34]. In healthcare, practitioners can rely on voice-guided teleoperation for precision in surgical robotics or patient care [38,40]. Additionally, the system's versatility can potentially be extended to fields like hazardous environment manipulation, educational robotics, and industrial automation, offering transformative possibilities for both expert and non-expert users [63].

The main contributions of this work are listed as follows:

– Propose a novel and pioneering controller-free robotic teleoperation system leveraging AR and LLMs.

- Integrate LLM-driven voice command functionality within an OST HMD-based AR robotic teleoperation framework.
- Develop a direct and effective AR-driven robotic puppeteering system that enhances user accessibility and intuitiveness.

2 Related Work

Extensive research has explored the integration of extended reality (XR)-encompassing virtual reality (VR), augmented reality (AR), and mixed reality (MR)-into robotics across various domains. In this section, we review the current landscape of VR, AR, and MR applications in robotics, with a particular focus on AR for human-robot interaction (HRI).

2.1 Virtual Reality in Robotics

The use of XR technologies to facilitate robotics research has gained significant traction over the past decades. VR, in particular, has been widely adopted for simulating robots due to its ability to provide precise and controlled virtual environments [27,31,43,53]. Burghardt et al. [7] developed a VR system with digital twins to capture human movements in a virtual environment, enabling industrial robots to replicate complex human actions that are otherwise difficult to automate. Similarly, Togias et al. [47] explored a VR-based teleoperation method for reprogramming industrial robots, enabling flexible and remote process design, which facilitates hybrid human-robot collaboration in dynamic production environments.

However, VR's integration with robotics has been particularly prominent in healthcare, especially for rehabilitation applications [1,6,49,54]. Over three decades ago, Ojha [36] recognized VR's potential in medicine and surgery, proposing its use for evaluating impaired hand performance, monitoring patient progress, and optimizing therapy strategies. Mirelman et al. [33] demonstrated that coupling robotic devices with VR can enhance post-stroke recovery by improving walking ability. Likewise, Frisoli et al. [16] developed a robotic exoskeleton integrated with VR for arm rehabilitation, showcasing its benefits in motor function recovery.

2.2 Augmented Reality for Human-Robot Interaction

The application of AR (often used interchangeably with MR) in HRI and robotic teleoperation [5,30,44,66] has become an established field. In recent years, AR has gained traction for robotic teleoperation due to its enhanced visualization capabilities and rich information presentation [35,62,67]. For instance, Hedayati et al. [23] demonstrated how AR can improve robot teleoperation by providing intuitive visual feedback. Their user study on AR-based aerial robot interfaces highlighted significant usability and performance improvements over traditional

methods that divide user attention between monitoring the robot and its camera feed. Additionally, Hashimoto et al. [21] introduced *TouchMe*, a touch-based AR interface for intuitive robot teleoperation from a third-person perspective. Unlike traditional joystick controls, users could directly manipulate robot components by interacting with a live camera feed, significantly reducing the learning curve. In industrial settings, Michalos et al. [32] proposed an AR tool to support hybrid human-robot workplaces, offering real-time visualizations, instructional overlays, and production updates, thereby improving operator safety, acceptance, and workflow integration. For human-robot communication, Walker et al. [50] explored AR-based motion intent cues to enhance interaction with non-anthropomorphic robots. Their study demonstrated that AR cues improve task efficiency and influence user perception of robots as teammates.

Regarding AR-based robotic puppeteering, Van et al. [48] introduced the first optical see-through head-mounted display (HMD) AR system for robotic puppeteering, where users manually controlled a virtual robot through direct input. While this approach provided a foundation for AR-driven robot manipulation, it still required explicit manual control, limiting the fluidity and accessibility of the interaction. Our work builds upon and extends this concept by introducing a controller-free AR robotic puppeteering system, eliminating the need for physical controllers and enhancing usability. By integrating voice commands powered by LLMs, we enable a more natural and intuitive way for users to command and direct robot movements. This advancement not only improves interaction flexibility but also lowers the barrier for users with minimal experience, fostering a more seamless and immersive puppeteering experience in AR-driven robotic control.

3 The System Introduction

3.1 Overview

In terms of the equipment, this innovative system is built in the Meta Quest 3 HMD, which immerses users in a dynamic AR environment. The virtual robot arm, modeled after the real Franka robot arm[1], serves as the primary interface for interaction and control. As shown in Fig. 3, the physical robot operates in a practical setting, while following the identical trajectory of the virtual robot projected in the AR settings.

The proposed LLM-driven, voice-controlled, controller-free robotic puppeteering system follows a structured pipeline, as depicted in the figure. The system integrates user interaction, AR view, and real-time robotic puppeteer to create an intuitive and interactive teleoperation experience.

1. **User Interaction:** This component highlights how users engage with the system through either voice commands or controller-free mid-air interactions. Voice commands, powered by LLMs, are processed to generate contextual

[1] https://franka.de/.

movement instructions for the virtual robot. The controller-free hand interaction enables users to intuitively spawn the virtual robot at a desired location within the AR environment.
2. **AR View:** Within the AR environment, the virtual robot arm acts as an intermediary, serving as a digital "puppet" modeled after the physical robot. Users define movements in AR, which are then transmitted to the real robot to be replicated. This mechanism ensures that all actions finalized in AR are accurately mirrored in the physical space, providing a seamless and interactive puppeteering process.
3. **Real-time Robotic Puppeteering:** The physical robot executes tasks in real time, precisely replicating the movements defined by the virtual robot. In our current implementation, this process involves placing the robot from an initial position to a predefined target position. This final step completes the interaction loop, allowing users to visually confirm the execution of their voice commands within the AR view, while ensuring precise 'puppeteer' of the physical robot.

3.2 User Demonstration

As illustrated in Fig. 3, we conducted an initial demonstration with a small group of users from the authors' local network. While this experiment successfully validated the core functionality of our proposed system, a more comprehensive pilot study and user evaluation remain to be designed and implemented.

The middle sub-figure in Fig. 3 showcases the AR interface projected in front of the user. Upon launching the dedicated application, the virtual robot was spawned alongside the user's hand, following mid-air controller-free interactions. As depicted in the left-side sub-figures, the user issues voice commands via a connected microphone to control the virtual robot, eliminating the need for physical controllers. The entire process is achieved through a controller-free mechanism, enabling intuitive and hands-free interaction.

The robotic puppeteering is executed through the robot arm's end-effector, which moves from its original position to a predefined target location, represented by a color-coded area, as shown in the right-side sub-figures. Once the virtual robot is spawned via hand interaction, the user issues a voice command – "Blueberry, move to [color]" – prompting the virtual robot's end-effector to navigate toward the specified color. Simultaneously, the real robot mirrors this trajectory, precisely replicating the virtual robot's movement in the real environment. Figure 3 shows the example of a movement to the [orange] area.

4 Technic Setup of the System

This section provides a detailed description of the technological setup used in our proposed puppeteer system. The setup consists of three core modules: voice recognition, locally hosted LLMs, and the pipeline for handling transcriptions and commands.

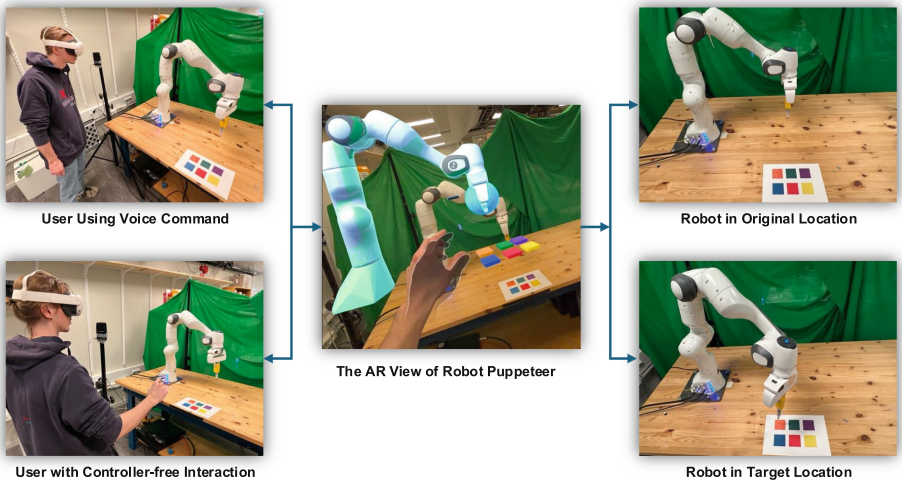


Fig. 3. A schematic overview of the proposed system with a user in action. The virtual Franka robot arm is rendered within the AR view (the middle one). Users can interact with the virtual robot through controller-free mid-air interaction by hands and voice commands signaled by a ZOOM H2N microphone (the left sub-figures), enabling precise manipulation of its movements (the right sub-figures). Simultaneously, the physical robot mirrors the virtual robot's trajectory in real-time, completing the 'puppeteer' process seamlessly.

4.1 Voice Recognition

For voice recognition and transcription, we utilized RealtimeSTT proposed by Kolja Beigel[2], a robust speech-to-text tool designed for real-time applications. The following configurations were applied:

- **Wakeword:** We employ the wakeword "Blueberry", which is one of the pre-trained options included in the RealtimeSTT framework. This wakeword triggers the voice recognition pipeline to start processing commands.
- **Speech Recognition Model:** The `tiny.en` model is used for lightweight and efficient transcription.
- **Wakeword Sensitivity:** The wake_words_sensitivity parameter is set to `0.9`, ensuring high sensitivity to the wakeword.
- **Microphone:** An external ZOOM H2N microphone is employed for audio input.

4.2 Locally Hosted LLMs

To process and validate commands, we run instances of locally hosted LLMs. This approach ensures low-latency interactions and data privacy. The setup includes:

[2] https://github.com/KoljaB/RealtimeSTT.

- **LLM Instances:** We use two instances of Llama 3.2 1B-Instruct Q6 models. These instances are hosted locally to maintain responsiveness and to avoid reliance on external servers.
- **Serving Framework:** The models are deployed using the `Llamafile`[3] framework, which provides an efficient and scalable mechanism for hosting multiple LLM instances on a local server.

4.3 Pipeline for Handling Transcriptions

Once the text is transcribed, it triggers multiple stages of validation and command handling, as detailed below:

1. **Mode Identification:** The system first checks the current operational mode. If the mode is set to "puppeteer", the pipeline proceeds to look for specific "move-commands". Other modes may invoke different processing logic.
2. **Command Validation:** Upon identifying potential commands, the system performs the following checks:
   (a) **Quick Validation:** The text is analyzed to determine if it contains the phrase "move to" followed by a predefined color (e.g., red, blue, orange, yellow, purple, green, or black (for tutorial only)).
      - If the command is valid ('yes'), the system returns a JSON object with the target color and proceeds to the next step.
      - If the command is invalid ('no'), the system proceeds to the previous validation stage.

   (b) **LLM Validation:** Using a round-robin approach, the system gets the current prompt for further validation. The LLM performs the following assessments:
      - If the command is identified as a valid "move-command" without any ambiguity, the system returns a JSON object with the target color and proceeds to the next step.
      - If the command is clearly not a "move-command", the system discards it and return to the previous step.
      - If the command is uncertain, the system marks the command for potential re-validation, which can be configured to custom rules.

   (c) **Command Execution:** The validated JSON object is sent to the OSC client via the UDP protocol.

5 Discussion and Future Work

In this section, we revisit our proposed system, reflect on its contributions, discuss its current limitations, and outline potential future work.

[3] https://builders.mozilla.org/project/llamafile/.

5.1 Insights

This work advances AR-based robotic teleoperation by introducing a controller-free, voice-driven interaction system powered by LLMs, significantly lowering barriers to robotic control. By addressing key HCI challenges – such as accessibility, intuitiveness, and ease of use – this system fosters natural, human-centric collaboration with robots. Unlike conventional robotic control methods or previous robotic puppeteering systems that rely on physical controllers or complex programming, our system enables users to engage with robots using context-aware voice commands and mid-air interactions. The AR interface improves usability by providing real-time visual feedback, allowing users to validate robotic movements instantly. Leveraging the Meta Quest 3 HMD, this system delivers an immersive robotic puppeteering experience, fostering greater confidence, precision, and engagement in robotic manipulation. This approach significantly reduces the learning curve for non-experts, making robotic teleoperation more intuitive, fluid, and accessible across various skill levels. By bridging advancements in NLP, AR, and robotics, this work lays the foundation for a more intuitive, flexible, and inclusive approach to robotic teleoperation.

A key strength of our system is its ability to decouple direct physical interaction from real-world robotic control, allowing users to manipulate a virtual robot while maintaining full situational awareness of the real robot. This approach minimizes risks associated with direct robotic operation, offering a safer and more intuitive method for teleoperation, particularly in high-stakes and industrial environments. By leveraging AR visualization and voice-driven commands, users can experiment with robot control in a low-risk virtual space, ensuring precision and confidence before executing real-world actions.

5.2 Potential Practical Use

The versatility of this system extends across multiple domains, offering practical solutions in industry, healthcare, hazardous environments, and education.

In industrial automation, it can streamline direct hands-on robotic control, enabling tasks such as assembly, maintenance, and quality inspection to be performed more efficiently and with less reliance on specialized training. For instance, in automotive manufacturing, workers could verbally instruct robotic arms to assemble components while monitoring their actions in real-time AR, reducing both training time and human error. Similarly, in warehouse logistics, workers could use gesture-based spawning and voice commands to remotely guide robotic pickers or conveyors, improving operational efficiency while reducing physical strain and workplace injuries. By minimizing direct physical engagement with heavy or hazardous machinery, this approach not only enhances productivity and accuracy but also significantly improves worker safety.

In healthcare, this technology has the potential to revolutionize assistive robotics, allowing individuals with limited mobility to interact with robotic systems through natural speech and AR visualization. This could be particularly impactful for stroke rehabilitation, where patients could guide a robotic

exoskeleton simply by speaking commands, helping them regain motor function without requiring extensive physical effort. Similarly, in elderly care, voice-controlled robotic assistants could help with daily activities, such as fetching objects, adjusting room settings, or even assisting with physical therapy, all while allowing caregivers to monitor progress through an AR interface.

For hazardous environments, such as disaster response, nuclear facility management, and deep-sea exploration, voice-driven teleoperation eliminates the necessity for human presence in high-risk conditions. Emergency responders could, for example, remotely control robotic search-and-rescue units inside collapsed buildings while receiving live AR feedback on environmental hazards. In nuclear decommissioning, engineers could manipulate robotic arms to safely handle radioactive materials without exposure, using AR overlays to guide their precision. Similarly, in deep-sea exploration, researchers could issue high-level verbal commands to autonomous underwater vehicles, allowing them to conduct environmental monitoring, specimen collection, or infrastructure inspections without requiring expert-level piloting skills. By allowing operators to control robots from a safe distance, this system significantly enhances worker safety and operational effectiveness in some of the world's most challenging environments.

Additionally, in education and research, our system provides an engaging and interactive platform for teaching robotic control principles without requiring coding or technical expertise. This can democratize access to robotics education, allowing students, educators, and researchers to experiment with robotic programming in an intuitive way. In practical classrooms, for example, students could manipulate virtual robots in AR, learning about kinematics, motion planning, and AI-driven control through hands-on interaction. Researchers in HRI could use this system to study intuitive command structures, user engagement, and learning curves, helping refine future robotic systems. By making robotics more accessible, this technology has the potential to foster innovation, inspire future generations, and accelerate advancements in HRI.

By enabling safe, efficient, and intuitive human-robot collaboration, this system bridges cutting-edge AR, NLP, and robotics technologies into a unified framework. Its broad applicability suggests that controller-free, voice-driven robotic control could play a pivotal role in shaping the future of automation, assistive technology, and intelligent robotics across industries.

5.3 Limitations

Despite the novel contributions, our system still has several limitations that must be addressed.

First, voice recognition accuracy remains an area for improvement. During user demonstrations, we observed that in some cases, the system misinterpreted voice commands, especially under noisy conditions or when users had diverse accents or speech patterns. The effectiveness of the LLM-driven voice processing is dependent on further refinement to ensure more reliable and responsive recognition. These observed issues can also be addressed by using larger models for the voice recognition by adding more computational power.

Second, while our controller-free interaction allows users to spawn the virtual robot via mid-air interactions through hand tracking parameters, it does not yet support more advanced hand gestures for direct manipulation for users but only supports developers for hand gesture customization. The lack of robust gesture-based controls limits the system's interactivity, which could otherwise enhance the overall user experience and precision of robotic control.

Third, our robotic puppeteering implementation is currently limited to moving or navigating the robot arm through a single trajectory with basic motion planning. In real-world applications, robotic tasks often involve complex, multi-step motions that require more precise motion planning and more sophisticated voice command processing. Expanding the system to support dynamic, real-time trajectory planning is crucial for practical deployment in industrial and research settings.

Lastly, although we conducted initial demonstrations with a small number of users, we have not yet performed a comprehensive user study with structured task design and systematic user evaluation. The lack of a formal user study limits our ability to quantitatively assess usability, efficiency, and user experience, which are essential for determining the system's real-world practicality.

5.4 Future Work

To address these limitations and enhance the system's capabilities, several aspects of future work are planned as follows.

First, we aim to improve voice recognition by integrating advanced speech models that are more robust to different accents, speech tones, and environmental noise. By using more powerful large models, we can make the voice command interpretation loop more inclusive and reliable.

Second, we will expand gesture-based interactions beyond virtual robot spawning. By incorporating natural hand gestures for robotic control, users will be able to manipulate the virtual robot more precisely, providing a more fluid and interactive experience. This multimodal interaction will allow for a combination of voice and hand gestures, leading to more seamless robotic teleoperation.

Third, we plan to increase the complexity of robotic motion by incorporating more sophisticated voice commands that facilitate dynamic, context-aware motion planning, while also accounting for the robot's degrees of freedom to enable more precise and adaptable control. This will allow robots to execute multi-step, adaptive movements, making them suitable for more complex tasks in industrial and assistive applications.

Finally, we will conduct a comprehensive user study to rigorously evaluate the system's usability, performance, and user experience. This study will involve a diverse group of participants across different expertise levels to assess the system's effectiveness, intuitiveness, and practical applicability. Through user feedback, we will refine the interaction design, responsiveness, and overall workflow to ensure a more seamless and efficient robotic teleoperation system.

By addressing these areas, we aim to push the boundaries of AR-driven robotic control, making it more practical, intelligent, safe, and accessible for a broad range of users across different industries.

6 Conclusion

This work presents a controller-free, LLM-driven voice-commanded AR robotic teleoperation system, marking a significant step toward more intuitive, accessible, and immersive HRI. By integrating AR and NLP, our approach eliminates the constraints of traditional input devices, enabling fluid and natural communication between humans and robots. This advancement lowers the barrier to robotic control, making it more inclusive for users across different skill levels and application domains.

Beyond enhancing accessibility and usability, this system establishes a versatile foundation for future developments in robotic teleoperation and interactive automation. Its potential extends to adaptive learning for task-specific customization, multimodal input integration, and collaborative robotics, where robots could intelligently respond to user commands in more dynamic and complex environments. Additionally, its ability to blend the physical and digital worlds through AR-driven visualization opens up new opportunities for safety, training, and real-time robotic supervision in industrial, medical, and other research settings.

Acknowledgment. This work was supported by the Swedish Foundation for Strategic Research (SSF) grant FUS21-0067.

References

1. Albani, J.M., Lee, D.I.: Virtual reality-assisted robotic surgery simulation. J. Endourol. **21**(3), 285–287 (2007)
2. Arevalo Arboleda, S., Rücker, F., Dierks, T., Gerken, J.: Assisting manipulation and grasping in robot teleoperation with augmented reality visual cues. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, pp. 1–14 (2021)
3. Baytas, M.A., Çay, D., Zhang, Y., Obaid, M., Yantaç, A.E., Fjeld, M.: The design of social drones: a review of studies on autonomous flyers in inhabited environments. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–13 (2019)
4. Breazeal, C., Dautenhahn, K., Kanda, T.: Social Robotics. Springer Handbook of Robotics, pp. 1935–1972 (2016)
5. Brizzi, F., Peppoloni, L., Graziano, A., Di Stefano, E., Avizzano, C.A., Ruffaldi, E.: Effects of augmented reality on the performance of teleoperated industrial assembly tasks in a robotic embodiment. IEEE Trans. Hum. Mach. Syst. **48**(2), 197–206 (2017)
6. Brütsch, K., et al.: Virtual reality for enhancement of robot-assisted gait training in children with central gait disorders. J. Rehabil. Med. **43**(6), 493–499 (2011)

7. Burghardt, A., Szybicki, D., Gierlak, P., Kurc, K., Pietruś, P., Cygan, R.: Programming of industrial robots using virtual reality and digital twins. Appl. Sci. **10**(2), 486 (2020)
8. Chan, W.P., et al.: Design and evaluation of an augmented reality head-mounted display interface for human robot teams collaborating in physically shared manufacturing tasks. ACM Trans. Hum. Robot Interact. (THRI) **11**(3), 1–19 (2022)
9. Chen, L., Lei, Y., Jin, S., Zhang, Y., Zhang, L.: Rlingua: improving reinforcement learning sample efficiency in robotic manipulations with large language models. IEEE Robot. Autom. Lett. (2024)
10. Chu, K., Zhao, X., Weber, C., Li, M., Wermter, S.: Accelerating reinforcement learning of robotic manipulations via feedback from large language models. arXiv preprint arXiv:2311.02379 (2023)
11. Cresswell, K., Cunningham-Burley, S., Sheikh, A.: Health care robotics: qualitative exploration of key challenges and future directions. J. Med. Internet Res. **20**(7), e10410 (2018)
12. De Franco, A., Lamon, E., Balatti, P., De Momi, E., Ajoudani, A.: An intuitive augmented reality interface for task scheduling, monitoring, and work performance improvement in human-robot collaboration. In: 2019 IEEE International Work Conference on Bioinspired Intelligence (IWOBI), pp. 75–80. IEEE (2019)
13. De Pace, F., Manuri, F., Sanna, A., Fornaro, C.: A systematic review of augmented reality interfaces for collaborative industrial robots. Comput. Ind. Eng. **149**, 106806 (2020)
14. Dzedzickis, A., Subačiūtė-Žemaitienė, J., Šutinys, E., Samukaitė-Bubnienė, U., Bučinskas, V.: Advanced applications of industrial robotics: new trends and possibilities. Appl. Sci. **12**(1), 135 (2021)
15. Fang, H., Ong, S., Nee, A.: Interactive robot trajectory planning and simulation using augmented reality. Robot. Comput. Integr. Manuf. **28**(2), 227–237 (2012)
16. Frisoli, A., et al.: Arm rehabilitation with a robotic exoskeleleton in virtual reality. In: 2007 IEEE 10th International Conference on Rehabilitation Robotics, pp. 631–642. IEEE (2007)
17. Gao, J., et al.: Physically grounded vision-language models for robotic manipulation. In: 2024 IEEE International Conference on Robotics and Automation (ICRA), pp. 12462–12469. IEEE (2024)
18. Ghiringhelli, F., Guzzi, J., Di Caro, G.A., Caglioti, V., Gambardella, L.M., Giusti, A.: Interactive augmented reality for understanding and analyzing multi-robot systems. In: 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1195–1201. IEEE (2014)
19. Goel, R., Gupta, P.: Robotics and Industry 4.0. A Roadmap to Industry 4.0: Smart Production, Sharp Business and Sustainable Development, pp. 157–169 (2020)
20. Green, S.A., Billinghurst, M., Chen, X., Chase, J.G.: Human-robot collaboration: a literature review and augmented reality approach in design. Int. J. Adv. Rob. Syst. **5**(1), 1 (2008)
21. Hashimoto, S., Ishida, A., Inami, M., Igarashi, T.: Touchme: an augmented reality based remote robot manipulation. In: The 21st International Conference on Artificial Reality and Telexistence, Proceedings of ICAT2011, vol. 2 (2011)
22. He, W., Li, Z., Chen, C.P.: A survey of human-centered intelligent robots: issues and challenges. IEEE/CAA J. Automatica Sinica **4**(4), 602–609 (2017)
23. Hedayati, H., Walker, M., Szafir, D.: Improving collocated robot teleoperation with augmented reality. In: Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, pp. 78–86 (2018)

24. Javaid, M., Haleem, A., Singh, R.P., Suman, R.: Substantial capabilities of robotics in enhancing industry 4.0 implementation. Cogn. Robot. **1**, 58–75 (2021)
25. Kim, C.Y., Lee, C.P., Mutlu, B.: Understanding large-language model (LLM)-powered human-robot interaction. In: Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction, pp. 371–380 (2024)
26. Kyjanek, O., Al Bahar, B., Vasey, L., Wannemacher, B., Menges, A.: Implementation of an augmented reality AR workflow for human robot collaboration in timber prefabrication. In: Proceedings of the 36th International Symposium on Automation and Robotics in Construction, ISARC, pp. 1223–1230. International Association for Automation and Robotics in Construction (IAARC) (2019)
27. Lee, J.Y., Mucksavage, P., Kerbl, D.C., Huynh, V.B., Etafy, M., McDougall, E.M.: Validation study of a virtual reality robotic simulator-role as an assessment tool? J. Urol. **187**(3), 998–1002 (2012)
28. Lee, Y.K., Jung, Y., Kang, G., Hahn, S.: Developing social robots with empathetic non-verbal cues using large language models. arXiv preprint arXiv:2308.16529 (2023)
29. Li, X., et al.: ManipLLM: embodied multimodal large language model for object-centric robotic manipulation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 18061–18070 (2024)
30. Marin, R., Sanz, P.J., Sánchez, J.S.: A very high level interface to teleoperate a robot via web including augmented reality. In: Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No. 02CH37292), vol. 3, pp. 2725–2730. IEEE (2002)
31. Matsas, E., Vosniakos, G.C.: Design of a virtual reality training system for human-robot collaboration in manufacturing tasks. Int. J. Interact. Des. Manuf. (IJIDeM) **11**, 139–153 (2017)
32. Michalos, G., Karagiannis, P., Makris, S., Tokçalar, Ö., Chryssolouris, G.: Augmented reality (AR) applications for supporting human-robot interactive cooperation. Procedia CIRP **41**, 370–375 (2016)
33. Mirelman, A., Bonato, P., Deutsch, J.E.: Effects of training with a robot-virtual reality system compared with a robot alone on the gait of individuals after stroke. Stroke **40**(1), 169–174 (2009)
34. Norberto Pires, J.: Robot-by-voice: experiments on commanding an industrial robot using the human voice. Ind. Robot Int. J. **32**(6), 505–511 (2005)
35. Nowak, A., Zhang, Y., Romanowski, A., Fjeld, M.: Augmented reality with industrial process tomography: to support complex data analysis in 3D space. In: Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers, pp. 56–58 (2021)
36. Ojha, A.K.: An application of virtual reality in rehabilitation. In: Proceedings of SOUTHEASTCON'94, pp. 4–6. IEEE (1994)
37. Pan, Y., Chen, C., Li, D., Zhao, Z., Hong, J.: Augmented reality-based robot teleoperation system using RGB-D imaging and attitude teaching device. Robot. Comput. Integr. Manuf. **71**, 102167 (2021)
38. Pulikottil, T.B., Caimmi, M., D'Angelo, M.G., Biffi, E., Pellegrinelli, S., Tosatti, L.M.: A voice control system for assistive robotic arms: preliminary usability tests on patients. In: 2018 7th IEEE International Conference on Biomedical Robotics and Biomechatronics (Biorob), pp. 167–172. IEEE (2018)
39. Riek, L.D.: Healthcare robotics. Commun. ACM **60**(11), 68–78 (2017)
40. Rogowski, A.: Scenario-based programming of voice-controlled medical robotic systems. Sensors **22**(23), 9520 (2022)

41. Sahin, M., Subramanian, K., Sahin, F.: Using augmented reality to enhance worker situational awareness in human robot interaction. In: 2024 IEEE Conference on Telepresence, pp. 217–224. IEEE (2024)
42. Sarker, S., Jamal, L., Ahmed, S.F., Irtisam, N.: Robotics and artificial intelligence in healthcare during COVID-19 pandemic: a systematic review. Robot. Auton. Syst. **146**, 103902 (2021)
43. Sethi, A.S., Peine, W.J., Mohammadi, Y., Sundaram, C.P.: Validation of a novel virtual reality robotic simulator. J. Endourol. **23**(3), 503–508 (2009)
44. Solanes, J.E., Muñoz, A., Gracia, L., Martí, A., Girbés-Juan, V., Tornero, J.: Teleoperation of industrial robot manipulators based on augmented reality. Int. J. Adv. Manuf. Technol. **111**, 1077–1097 (2020)
45. Song, D., Liang, J., Payandeh, A., Raj, A.H., Xiao, X., Manocha, D.: VLM-social-nav: socially aware robot navigation through scoring using vision-language models. IEEE Robot. Autom. Lett. (2024)
46. Tapus, A., Mataric, M.J., Scassellati, B.: Socially assistive robotics [grand challenges of robotics]. IEEE Robot. Autom. Mag. **14**(1), 35–42 (2007)
47. Togias, T., Gkournelos, C., Angelakis, P., Michalos, G., Makris, S.: Virtual reality environment for industrial robot control and path design. Procedia CIRP **100**, 133–138 (2021)
48. Van Haastregt, J., Welle, M.C., Zhang, Y., Kragic, D.: Puppeteer your robot: augmented reality leader-follower teleoperation. In: 2024 IEEE-RAS 23rd International Conference on Humanoid Robots (Humanoids), pp. 1019–1026. IEEE (2024)
49. Wade, E., Winstein, C.J.: Virtual reality and robotics for stroke rehabilitation: where do we go from here? Top. Stroke Rehabil. **18**(6), 685–700 (2011)
50. Walker, M., Hedayati, H., Lee, J., Szafir, D.: Communicating robot motion intent with augmented reality. In: Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, pp. 316–324 (2018)
51. Walker, M.E., Hedayati, H., Szafir, D.: Robot teleoperation with augmented reality virtual surrogates. In: 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 202–210. IEEE (2019)
52. Wallace, B., et al.: Imitation or innovation? translating features of expressive motion from humans to robots. In: Proceedings of the 12th International Conference on Human-Agent Interaction, pp. 296–304 (2024)
53. Whitney, D., Rosen, E., Phillips, E., Konidaris, G., Tellex, S.: Comparing robot grasping teleoperation across desktop and virtual reality with ROS reality. In: Robotics Research: The 18th International Symposium ISRR, pp. 335–350. Springer (2019)
54. Xia, H., et al.: Shaping high-performance wearable robots for human motor and sensory reconstruction and enhancement. Nat. Commun. **15**(1), 1760 (2024)
55. Zeng, F., Gan, W., Wang, Y., Liu, N., Yu, P.S.: Large language models for robotics: a survey. arXiv preprint arXiv:2311.07226 (2023)
56. Zhang, C., Chen, J., Li, J., Peng, Y., Mao, Z.: Large language models for human-robot interaction: a review. Biomimetic Intell. Robot. 100131 (2023)
57. Zhang, Y., et al.: Human-centered AI technologies in human-robot interaction for social settings. In: Proceedings of the International Conference on Mobile and Ubiquitous Multimedia, pp. 501–505 (2024)
58. Zhang, Y., Ma, Y., Kragic, D.: Vision beyond boundaries: an initial design space of domain-specific large vision models in human-robot interaction. In: Adjunct Proceedings of the 26th International Conference on Mobile Human-Computer Interaction, pp. 1–8 (2024)

59. Zhang, Y., Nowak, A., Rao, G., Romanowski, A., Fjeld, M.: Is industrial tomography ready for augmented reality? a need-finding study of how augmented reality can be adopted by industrial tomography experts. In: International Conference on Human-Computer Interaction, pp. 523–535. Springer (2023)
60. Zhang, Y., Nowak, A., Romanowski, A., Fjeld, M.: On-site or remote working?: an initial solution on how COVID-19 pandemic may impact augmented reality users. In: Proceedings of the 2022 International Conference on Advanced Visual Interfaces, pp. 1–3 (2022)
61. Zhang, Y., Nowak, A., Xuan, Y., Romanowski, A., Fjeld, M.: See or hear? exploring the effect of visual/audio hints and gaze-assisted instant post-task feedback for visual search tasks in AR. In: 2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 1113–1122. IEEE (2023)
62. Zhang, Y., Omrani, A., Yadav, R., Fjeld, M.: Supporting visualization analysis in industrial process tomography by using augmented reality-a case study of an industrial microwave drying system. Sensors **21**(19), 6515 (2021)
63. Zhang, Y., et al.: Mind meets robots: a review of EEG-based brain-robot interaction systems. arXiv preprint arXiv:2403.06186 (2024)
64. Zhang, Y., Vasco, M., Björkman, M., Kragic, D.: Will you participate? exploring the potential of robotics competitions on human-centric topics. In: International Conference on Human-Computer Interaction, pp. 240–255. Springer (2024)
65. Zhang, Y., Xuan, Y., Yadav, R., Omrani, A., Fjeld, M.: Playing with data: an augmented reality approach to interact with visualizations of industrial process tomography. In: IFIP Conference on Human-Computer Interaction, pp. 123–144. Springer (2023)
66. Zhao, Z., Huang, P., Lu, Z., Liu, Z.: Augmented reality for enhancing tele-robotic system with force feedback. Robot. Auton. Syst. **96**, 93–101 (2017)
67. Zhong, X., Liu, P., Georganas, N.D., Boulanger, P.: Designing a vision-based collaborative augmented reality application for industrial training. IT Inf. Technol. **45**(1), 7–19 (2003)

Social Media in Learning and Education

Redimensioning Visible Learning and Teaching in the Dynamics of a New Reality

Rudina Alimerko(✉) and Deniz Çupi Xhoga

Faculty of Humanities, "Ismail Qemali" University, Vlorë, Albania
{rudina.alimerko,deniz.cupi}@univlora.edu.al

Abstract. The complexity of globalization, cutting-edge technology, and societal change are all reflected in the massive challenges faced by the education sector in the twenty-first century. In order to shape citizens who can contribute to a just and sustainable society, education policies and the role of teachers should be reformed to adjust to the dynamics of the new reality. This study aims to examine the effectiveness of the teaching approach based on John Hattie's concept of visible learning, in order to enhance learning outcomes and the overall education experience. The research methodology relies on a qualitative descriptive approach, which combines theoretical and practical analysis. Specifically, a case study (practical case study) has been analysed during the mandatory preparatory teacher training. Results indicate that teachers implementing Hattie's approach share successful experiences and methods. The findings reveal that the effect of this method is evident in multiple significant aspects, concerning readers' engagement with the content of the novel and the way this novel affects their personal and intellectual growth. With the assistance of artificial intelligence (AI) and the application of generative AI (GenAI), which pertains to technology in teaching and learning, in order to provide students and teachers with options on how to explore accessible alternatives as needed, theoretical and practical analyses of the experiences acquired can be performed in a more comprehensive and inclusive manner, which contributes to presenting a fresh aspect of effectiveness in teaching and lifelong learning.

Keywords: redimensioning · visible learning · artificial intelligence · critical reflection

1 Introduction

In a world marked by continuous conflicts, economic crises, pandemics, and technological transformation, the education sector needs to adapt its methodologies and practices to address these new circumstances. "The backdrop to 21st-century education is our endangered environment. Growing populations, resource depletion and climate change compel all of us to think about sustainability and the needs of future generations" [21]. It is essential to offer education policies and practices in establishing a sustainable and inclusive education system, which can handle the persistent uncertainties and challenges of the contemporary world. To rise to the challenges of our time, a move towards more

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 115–124, 2025.
https://doi.org/10.1007/978-3-031-93539-8_8

inclusive education is imperative Rethinking the future of education is all the more important following the Covid-19 pandemic, which further widened and put a spotlight on inequalities. Failure to act will hinder the progress of societies", noted the Director-General of UNESCO [29]. **Continuous training of teachers** on technology and innovative pedagogies is necessary to meet the challenges of a new educational reality. Educational institutions should provide opportunities for continuous professional development. **The development of flexible and appropriate** curricula that integrate technology and incorporate active learning approaches is an essential step to prepare learners for an increasingly digitalized and globalized world. Information and Communication Technology (ICT) has transformed every aspect and area of our lives [1]. At European level, these competences are included in a competence framework for teachers [12]. Likewise, UNESCO has drafted the ICT Competency Framework for Teachers (ICT CFT, Version 3), as a document to guide the development of teachers in the use of ICT. This framework aims to support countries in policy development and professional capacity building in the field of ICT in education [28]. In short, the role of teachers in education will be transformed permanently. We can hope for a better future in education as we move on towards the path of digitalization [17]. In order to shape citizens who can contribute to a just and sustainable society, education policies and the role of teachers should be reformed to adjust to the dynamics of the new reality. This will ensure that the requirements of a globalized world are harmonized with the maintenance of essential human and cultural values. This study aimed to explore the role of teachers in a new reality aiming to shape global citizens able to adjust to the ever-changing dynamics of a new reality.

2 Literature Review

Teachers ought to be ready to acknowledge the variety of students and to implement individualized teaching approaches. **Education for the future** involves the development of critical, creative and teamwork competencies. This necessitates that teachers inspire learners to consider perspectives beyond conventional boundaries and cultivate abilities that are essential for an evolving job market and society. Universities should encourage students on how to think and not what to think. To think critically is a process, and its acquisition takes time. Within this process, university teachers must be aware of the need for a transformation to adjust their teaching methodologies to the learning of their students, as their way of thinking may generate fixed routines in their teaching practices [14].

John Hattie's "*Visible Learning*" theoretical framework has exerted a profound influence on how teacher professional development is conceived and implemented in the context of contemporary education. At the core of this approach lies the focus on the effectiveness of teaching methods and the ability of teachers to have a measurable impact on students' performance and learning. This evidence-based model and empirical research fosters a culture of continuous reflection and improvement, where teachers are encouraged to identify practices that have the most significant positive effect on learning. Through the use of clear measurements and data from research, teachers get skilled in evaluating and reorganising their pedagogical strategies, thereby improving their professional skills. Hattie emphasizes the importance of developing a clear vision on what

constitutes an effective teaching, including the acquisition of appropriate tools to analyse and adapt their practices based on these values. He has also identified a range of factors that improve students'performance, including: computer tutoring that provides immediate feedback, particularly when using artificial intelligence [8]. It is believed that attempts have been undertaken in this paper to rethink teaching by employing creative approaches to meet the needs of the current times. As educators, it is our duty to reflect on being flexible to the visible learning process, which aims to clarify the thinking processes of learners during their educational experience. Considering the significance of learning, this research will be expanded to encompass additional training and empirical studies. This combined methodology offers an extensive summary of how teaching and learning are being redimensioned to meet the requirements of a new reality, blending hands-on experience and theoretical reflection, to produce a thorough and comprehensible analysis, to examine the challenges of education in the current era, and to propose tactics for its sustainable development.

We think that in more in-depth studies, John Hattie's approach to visible learning can be improved by using artificial intelligence to collect and analyse real-time data on student achievements and challenges when it comes to targeted and effective focus groups. It is therefore important to understand the impact of AI in the context of future learning and education, instead of in current systems of education and forms of learning. The analysis of the impact of AI on teaching will, therefore, be inherently linked to foresight-oriented work on the future of learning [10].

With the help of artificial intelligence, theoretical and practical analyses of teachers' experiences can be conducted in a more detailed and thorough manner, which contributes to providing a new level of effectiveness in teaching and learning. Every day, everywhere, AI is gaining popularity. An increasing number of students are specializing in it, preparing for careers that are either within it or closely related to it, thus it is essential to acknowledge it and incorporate it into the redimensioning of the learning process [11]. According to the researchers, the intersection of artificial intelligence (AI) and literature now embarks on a new realm of creating plots to develop stories and characters; how the landscape of storytelling has changed with the invention of new and advanced AI technologies, changing the nature of engaging tales as literature has otherwise always been carried out [13]. It is worth mentioning the use of generative AI (GenAI) which refers to technology in teaching and learning, to give students alternative options to perform a task, or to create new content such as text, images, codes, or ideas on how to explore accessible alternatives as needed [27]. GenAI can also spark creativity in students. [4] point out that the demands generated by AI offer different, unpredictable ideas, driving thought and innovative solutions. It is important to understand the factors that influence the increase of engagement and motivation of students in collaborative learning environments supported by AI technology to reflect on innovative strategies and the impact of these tools on educational, discipline and subject related outcomes.

The goal is to understand how the integration of AI and game elements can enhance the learning experience and promote collaboration between students, as well as to discuss about the role of GenAI gamification tools in higher education and its potential to transform the traditional educational system while upholding academic integrity [9]. In this context, the use of both general GenAI platforms and specific educational GenAI

tools should be designed to enhance teachers' understanding of their subject as well as their knowledge on teaching methodologies, including through teacher-AI co-designing of lesson plans, course packages, or entire curricula [30].

3 Study Methodology

This study relies on a qualitative approach, which combines theoretical and practical analysis, in the view of redimensioning visible learning and teaching in the dynamics of a new reality. This study aims to examine the effectiveness of the teaching approach based on John Hattie's concept of visible learning, in order to enhance learning outcomes and the overall education experience. At this stage, an in-depth analysis of the scientific literature has been made that addresses visible learning, contemporary pedagogies, the integration of technology into education, and the challenges of a new education reality.

Specifically, a case study (practical case study) has been analysed during the Compulsory Preparatory Teacher Training Program for Principals of Pre-University Education Institutions, with the participation of 67 teachers. This program is based on the aspects of teaching, learning processes, acquisition of the appropriate knowledge and skills for leadership, empowerment and motivation, including the encouragement of individual growth for the candidates aspiring to be principals. In this context, the role of teachers would be important and inclusive, so the focus will shift from acquiring knowledge to developing critical skills, such as critical thinking, problem solving, creativity, and collaboration. Hattie, in his extensive study on factors affecting learning, emphasizes the importance of clear interaction between teachers and students, promoting an approach where students see their achievements and where teachers become more aware of their impact on teaching. This approach is based on the use of data and reflection on teaching strategies, in order to maximize benefits for students.

Viewed from this perspective, during the training on the module "*Effective teaching and learning",* the implementation of *Visible Learning* was motivating for teachers because it created a friendly environment emphasizing how their reflection and self-assessment on teaching methods, helped to optimize the best pedagogical and didactic practices. It should be noted that one of the most important factors according to Hattie is to provide immediate and clear feedback, which helps students understand where they stand and what actions they can take to enhance their performance. The results, while addressing the most effective methods and techniques, demonstrate that teachers, in the implementation of Hattie's approach, share successful experiences and methods. This exchange helps to find solutions to specific challenges and to improve learning outcomes. It should be mentioned that during the sessions there were some difficulties regarding their role in the teaching process, for example: Teachers should transition from the conventional role of being the transmitter of information to a position where they serve more as mentors and who reflect on learning.

This change may require time and a more active approach to self-assessment on the part of teachers; Resistance to changes: Some teachers might experience discomfort when altering their conventional approaches and could be hesitant to adopt new strategies, such as consistently reflecting on their impact and employing effective feedback; Use of technology equipment: Technology can serve a crucial function to assist the feedback process and track the advancement of students. However, the lack of technological

skills of teachers, or the lack of equipment and infrastructure in some schools, can make it difficult to fully implement this approach.

These challenges show that, although Hattie's approach to "visible learning" has great benefits, it requires a profound transformation of teaching practices, requiring sufficient support, preparation and resources to achieve success in every context.

4 Outcomes

The practical case study approach allows for an in-depth exploration of the processes that occur in the real classroom context and can highlight the specific benefits and challenges arising from the implementation of this approach. We believe that in our case, this empirical approach has a beneficial impact on the professional growth of teachers, highlighting the significance of utilizing data to enhance teaching methods. Hattie analyses a range of studies related to the impact of teaching and pedagogical methods on student achievements, providing a clear framework for comprehending what is most effective for learning. Visible Learning presents research involving many millions of students and represents the largest ever collection of evidence-based research into what actually works in schools to improve learning. Areas covered include the influences of the student, home, school, curricula, teacher, and teaching strategies. A model of teaching and learning is developed based on the notion of visible teaching and visible learning [6].

A major message within the book is that what works best for students is similar to what works best for teachers. This includes an attention to setting challenging learning intentions, being clear about what success means, and an attention to learning strategies for developing conceptual understanding about what teachers and students know and understand.

To enhance comprehension of the effectiveness of Hattie's empirical methodology, qualitative data gathering techniques were employed, such as auditorium observations with primary education students and focus group discussions, to share experiences about the approach in the context of the "*Harry Potter"* series by J. K. Rowling. *The "Visible Learning"* approach, as coined by John Hattie, focuses on the importance of the learning process and raising students' awareness of their progress through feedback, reflection and self-assessment [7].

In the context of the *Harry Potter* series by J.K. Rowling, the impact of this approach on readers' development and their experience with the novel can be seen in several key aspects, related to the interaction of young readers with the content of the book and the way this novel affects their personal and intellectual development. Learning through literature can help you learn or gain productive experiences, skills or knowledge, as in our case with the magical world of Harry Potter [23]. Therefore, education should enable the power to continue research and learning endlessly [25].

A study published last year in Science found that reading literary fiction, as opposed to popular fiction or nonfiction, results in keener social perception and increased empathy which is being defined more or less as the ability to alternate between different perspectives on a particular person or situation. Literature with complex, developed themes and characters appears to let readers occupy or adopt perspectives they might otherwise not

consider; and it seems that Rowling might get at the beautiful, sobering mess of life in a way that could have a meaningful impact on our children's collective character [22].

Regarding the process of learning awareness (metacognition) in the novel "Harry Potter," students are exposed to the processes of learning and character development. Characters like Harry, Hermione, and Ron experience visible development through their experiences and lessons at Hogwarts School of Witchcraft and Wizardry. This approach can assist students and readers in enhancing their capacity to think about what they understand while reading and become more conscious of how they assimilate information. For example, Harry's experience with the Patronus, a technique he must acquire to defeat the Dementors, is an example of learning through continuous reflection and practice [19]. Students can relate to his endeavours and enhance their understanding of the learning journey, contemplating the individual efforts they put forth to refine their abilities.

Role of Feedback: Feedback is a key component of the approach [7]. In the novel "*Harry Potter*," teachers and mentors provide on-going feedback that aids the protagonists in enhancing their abilities. For example, Professor Lupin, Harry′s favourite teacher, a great friend of his parents, teaches Harry the Patronus charm to protect himself from the Dementors. Lupin provides the boy tools to create light and comprehend what is positive [2]. And Professor Dumbledore, offers safety and stability among children. He becomes somewhat of a guardian for Harry, albeit from afar [16]. Dumbledore seems quite content overseeing the education of young witches and wizards. Though he does not teach an actual class nor provide in-service lessons for teachers, he still succeeds in teaching both the students and faculty members of Hogwarts. By using his own demeanour as a model for others, Dumbledore strives to teach teachers and students alike about love. Dumbledore uses love to teach Harry about horcruxes, Voldemort's past, and Harry's own past, and these provide Harry with feedback regarding his advancement and ways for improvement [31]. This fosters a mutual relationship between the characters and their learning experiences, which reflects a teaching approach that makes progress visible. In our case, students and readers can benefit from this by experiencing the importance of feedback in enhancing their individual skills. By engaging with text and character experiences, they are able to contemplate the significance of feedback in daily life and cultivate a more aware and receptive attitude toward criticism and enhancement. Critical thinking is a capability being required by students in their personal and professional life. Thus, universities need to put forth their best efforts to incorporate it into their courses, programs, and syllabi [3].

Self-assessment and autonomous learning: One of the lessons Harry Potter encourages is the importance of self-assessment and autonomy in the learning process. Hermione Granger is an example of a student who uses self-assessment to evaluate her progress. She continually pursues further knowledge, studies more than the material being presented, and is self-critical in her quest for perfection. Readers in general, particularly young individuals, as well as students engaged in discussions and analyses, can gain from this and feel motivated to assume responsibility for their own learning, cultivating a mindset of self-assessment and self-management. They recognize that achieving academic success and personal growth necessitates effort and a proactive approach for on-going development.

Learning through experience and mistakes: Another important element of "visible learning" that can be applied to the novel "Harry Potter" is learning through experience and mistakes. Characters in the book series frequently gain insights from their experiences and the mistakes they commit throughout the journey. For example, Harry and his friends often make the wrong decisions, but through the reflection and feedback of their mentors, they learn to improve their behaviours and strategies. This teaches students that mistakes are not the end of learning, but a crucial part of the development process. The "visible learning" approach emphasizes the importance of mistakes in learning and that through reflection on them, improvement can be achieved.

Motivation and commitment for attaining goals: Characters in the novel "Harry Potter" often face challenges that require a high motivation and dedication to achieve their goals. This corresponds with Hattie's philosophy on the need to encourage students to be motivated and engaged in learning. Students who engage with the narrative of the characters may be motivated to emulate their actions in confronting challenges and striving for self-improvement [6]. This approach helps students and readers develop a positive outlook on difficulties and view their progress as something attainable through commitment and lessons learned. The Harry Potter series helps develop a rich environment for "visible learning", cultivating an understanding of learning, the significance of feedback, self-assessment and the experience of mistakes as elements of personal growth.

During the analysis students involved in this story encounter essential aspects that represent Hattie's methodology, assisting them in gaining greater insight into their learning processes and how to continually improve [5]. This empirical research focuses on teaching practices that are most effective for students, using measurable data to analyse teaching and learning success. In the context of book reading, this approach aims to maximize the benefits of reading in order to develop advanced understanding and analysis skills.

In this framework, visible learning includes three key aspects:

Clear learning objectives that require to link the reading of books to clear learning goals. This includes orienting towards critical reading skills, text analysis and the development of interpretative skills; so for example in Harry Potter and the Philosopher's Stone [18], the learning process within the concept of "visible learning", becomes effective and understandable, in the case where Professor Flitwick clearly defines at the beginning of the class the purpose of classroom learning, defining the steps, articulation, specific command (Wingardium Leviosa), and the precise movement of the magic wand.

The goal is clear: focus and motivation yield favourable and tangible outcomes; Provision of effective feedback. An essential aspect of Hattie's model is the provision of ongoing and specific feedback that helps students understand how they are progressing. In reading, this may include discussions about the book's content, reflections on the actions of the characters, or critical questions about events. Reading, as a primarily solitary activity, not only offers entertaining and educational content but also challenges us to envision life through another person's eyes, increasing our empathetic abilities [26]. Engaging students in self-assessment. Students are urged to be mindful of their advancement, contemplating what they have read and recognizing how this reading influences their personal and academic growth. Once students are interested in the stories, reading them not only inspires their creativity, but also teaches kids how to go about

reading a book. The books are filled with invented words to describe people, places, and spells.

Because so many words are invented, young readers also learn to discern meanings of words from the context they appear in, a valuable reading comprehension skill useful for the rest of their lives. Through read-alouds from their teachers, or reading on their own, students also learn to listen to the sound of words on a page, rather than just plodding through text skimming for information. Rowling excels at creating words whose sound connotes meaning. The skill of listening to the sounds of words on a page is essential for developing reading skills and creating a love of language. It even could help interest students in other languages: When Harry Potter is soon translated into Latin and Ancient Greek (Dowd, op-ed), students may be interested in these scholarly languages because they will have a modern cultural context in which to learn them [15]. A key premise is that for students to be able to improve, they must develop the capacity to monitor the quality of their own work during actual production [20]. Visible Learning in reading encourages teachers and students to focus on deep-thinking strategies and the deliberate use of books to advance their knowledge and skills.

5 Conclusion and Recommendations

The **use of technology** in education makes it possible to make teaching and learning more personalised and accessible. Teachers should develop technological competencies to integrate new education platforms and communication tools into the learning process. **Active interaction and engagement** are pivotal elements for visible learning. Students require a setting that actively inspires and engages them, providing them with the chance to share ideas, take part in conversations, and collaborate on innovative projects. **Differentiated instruction** according to students' different learning styles and needs is essential to support individual development and ensure everyone's success. **Support for teachers and students in terms of emotional well-being** during major changes is crucial. Education should not only address the academic aspects but also support the social and emotional well-being of individuals. **Encouraging collaborative work and learning through projects** is a strategy that helps increase student engagement and develop skills relevant to teamwork and problem-solving. **Continuous monitoring and improvement of teaching and learning** through evidence-based assessments will help adapt methods and practices to new educational needs and trends.

Considering that AI is currently a top priority on the policy agenda, it is quite simple to create high-level perspectives for the future that claim that AI is the next technological revolution. AI is now often called "the next electricity". It is therefore important that teachers, who often struggle with concrete demands of everyday teaching practice and new initiatives, will not be electrocuted by this new technology [10]. The literature reveals consensus on the importance of incorporating GenAI into higher education. Common themes like mentorship, personalised learning, creativity, emotional intelligence, and higher-order thinking highlight the persistent need to align human-centred educational practices with the capabilities of GenAI technologies [24].

References

1. ASCAP, Manual trajnimi Përdorimi i TIK-ut në procesin e mësimdhënie–nxënies. (Manual Use of ICT in the teaching-learning process), p. 4 access on web (2020) https://www.ascap.edu.al/wp
2. content/uploads/2020/10/FINAL-MANUALI-I-TRAJNIMIT-PER-ICT-16.03.2021–1.pdf
3. Bishqemi, A.: Histori e Letërsisë Botërore për Fëmijë. (History of World Children′s Literature), pp.314–315. Elbasan. ISBN 99943–704–7–2 (2006)
4. Bezanilla, M.J., Fernández-Nogueira, D., Poblete, M., Galindo-Domínguez, H.: Methodologies for teaching-learning critical thinking in higher education: the teacher's viw. Thinking Skills Creativity **33**, 100584 (2019). https://www.sciencedirect.com/science/article/abs/pii/S1871187
5. Chan, C.K., Hu, W.: Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. Int. J. Educ. Technol. High. Educ. **20**(1), 43 (2023)
6. Fisher, D., Frey, N., Hattie, J.: Visible learning for literacy, grades K-12: implementing the practices that work best to accelerate student learning. Thousand Oaks, California, Corwin/A SAGE Company (2016)
7. Hattie, J.: Visible learning: a synthesis of over 800 meta-analyses relating to achievement. Routledge (2009)
8. Hattie, J., Timperley, H.: The power of feedback. Rev. Educ. Res. **77**(1), 81–112 (2007). https://doi.org/10.3102/003465430298487
9. Hattie, J.: John hattie on the factors that influence learning in schools. social science space (2023) https://www.socialsciencespace.com/2023/03/john-hattie-on-the-factors-that-influence-learning-in-schools/
10. Hysaj, A., Khan, S.A., Hamam, D., Farouqa, G.: review of the factors that influence learners motivation in using ai-enabled gamification for collaborative learning. In: Zaphiris, P., Ioannou, A., Sottilare, R.A., Schwarz, J., Rauterberg, M. (eds.) HCI International 2024 – Late Breaking Papers. HCII 2024. Lecture Notes in Computer Science, vol. 15378. Springer, Cham (2025). https://doi.org/10.1007/978-3-031-76815-6_5
11. Tuomi, I.: The impact of artificial intelligence on learning, teaching, and education. Cabrera Giraldez, M., Vuorikari, R. and Punie, Y. (eds.), EUR 29442 EN, Publications Office of the European Union, p. 31, 32. Luxembourg (2018). ISBN 978–92–79–97257–7, https://doi.org/10.2760/12297, JRC113226
12. Kissinger, H.A., Schmidt, E., Huttenlocher, D.: Epoka e inteligjencës atificiale. (the age of artificial intelligence). Shtëpia Botuese Fan Noli, Fan Noli Publishing House, Tiranë (2024)
13. EU, European framework for digital competence of teachers, DigComEdu. (Korniza evropiane për kompetencën digjitale të mësuesve), access on web. https://ec.europa.eu/jrc/en/digcompedu
14. Kusmiatun, A., Efendi, D., Pansori, M.J.A., Judijanto, L., Sariasih, Y., Saddhono, K.: The power of AI in generating novel based and impactful character development for fiction story. In: 2024 4th International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), pp. 1555–1561. Greater Noida, India (2024). https://doi.org/10.1109/ICACITE60783.2024.10616700
15. Mahdi, O.A., Nasar, I.A., Almuslamani, H.A.: The role of using case studies method in improving students'critical thinking skills in higher education. International J. High. Educ. (2020) https://doi.org/10.5430/ijhe.v9n2p297
16. Pastoor, K.: Magic in the classroom. the controversial harry potter publisher by prized writing, access on web. https://prizedwriting.ucdavis.edu/magic-classroom-controversial-harry-potter?utm_source=chatgpt.com

17. Qafzezi, E.: Aspekte teorike dhe praktike në fushën e përkthimit të letërsisë për fëmijë. (Theoretical and practical aspects in the field of children′s literature translation), Tirane, p. 139 (2014)
18. Smile foundation. role of a teacher transformed in the era of digital learning (2023). https://www.smilefoundationindia.org/blog/role-of-a-teacher-transformed-in-the-era-of-digital-learning/
19. Rowling, K.J.: Harry Potter and the Philosopher′s Stone. (Harry Potter dhe Guri Filozofal) (1997)
20. Rowling, K.J.: Harry Potter and the Prisoner of Azgaban. Harry Potter dhe i Burgosuri i Azgabanit (1999)
21. Sadler, D.R.: Formative assessment and the design of instructional systems. Instr. Sci. **18**, 119–144 (1989). https://doi.org/10.1007/BF00117714
22. Schleicher, A.: world class: how to build a 21st-century school system (2018). https://doi.org/10.1787/9789264300002-en
23. Stetka, B.: Why everyone should read harry potter. tales of the young wizard instill empathy, a study finds (2014) access on web. https://www.scientificamerican.com/article/why-everyone-should-read-harry-potter/
24. Shyti, Sh., Kurrikula dhe Shkolla. Cikli Fillor. (Curriculum and School. Primary Cycle). Publisher by Institute of Curricula and Standards, Tirana, pp. 126–139 (2005)
25. Tillmanns, T., et al.: Mapping tomorrow's teaching and learning spaces: a systematic review on GenAI in higher education. Trends High. Educ. **4**(1), 2 (2025). https://doi.org/10.3390/higheredu4010002
26. Thomas, J.H.: Fuqia e Arsimit. (The Power of Education). Tiranë (2020)
27. The Reading Ranch.: The link between reading and emotional intelligence: a friendly exploration (2024) access on web. https://www.readingranch.com/the-link-between/
28. UC Berkeley. Undergraduate Education. Research, Teaching & Learning. AI in Teaching & Learning Overview (2024). https://rtl.berkeley.edu/ai-teaching-learning-overview
29. UNESCO, ICT competency framework for teachers, pp. 6 (2018), access on web. https://unesdoc.unesco.org/ark:/48223/pf0000265721_eng
30. UNESCO, Global Education Monitoring (GEM) Report: Inclusion and Education (2020) access on web. https://www.unesco.org/en/articles/global-education-monitoring-gem-report-2020
31. UNESCO, Guidance for generative AI in education and research. United National Educational, Scientific and Cultural Organization, pp. 31 (2023), access on web: https://unesdoc.unesco.org/ark:/48223/pf0000386693
32. Vaughn, M.E.: Keeping it real: teaching and learning in the harry potter series. Honors Theses. 1. p. 45, 55 (2011). https://aquila.usm.edu/honors_theses/1

Generative AI in Education: Exploring EAP Faculty Perspectives at a Multicultural UAE University

Asim Bashir[1](✉), Saima Nomaan[1], and Ajrina Hysaj[2]

[1] University of Birmingham Dubai, Dubai, UAE
{a.bashir.2,s.nomaan}@bham.ac.uk
[2] University of Wollongong in Dubai, Dubai, UAE
ajrinahysaj@uowdubai.ac.ae

Abstract. This study examines the perceptions and readiness of English for Academic Purposes (EAP) faculty at a multicultural university in the UAE to integrate Generative Artificial Intelligence (GenAI) into their teaching practices. Employing a qualitative case study approach, the research explores faculty perspectives through digital written interviews to understand the complex interplay between technological acceptance and pedagogical integration within a diverse educational setting. The findings reveal a spectrum of faculty attitudes towards GenAI, categorised into three main themes: Optimistic Integration, Cautious Acceptance, and Resistance and Scepticism. Faculty demonstrating Optimistic Integration perceive GenAI as a transformative tool that enhances educational delivery, personalises student learning, and automates routine tasks. Conversely, those with Cautious Acceptance recognise GenAI′s potential benefits but express concerns about ethical implications and the potential for technology to undermine academic integrity. The third group, exhibiting Resistance and Scepticism, focuses on the potential negative impacts of GenAI, including depersonalisation of education and loss of faculty jobs. The study underscores the need for robust faculty development programmes that address both the technological and pedagogical dimensions of GenAI. These programmes should promote understanding and skills necessary for the effective integration of GenAI, considering the institution′s multicultural and diverse linguistic backdrop. By aligning GenAI implementation with educational goals and ethical standards, the findings aim to guide policymakers, curriculum developers, and educational technology designers in fostering environments that leverage AI to enhance teaching effectiveness while respecting diverse academic needs.

Keywords: EAP faculty · Faculty perceptions · GenAI · Multicultural education

1 Introduction

Generative Artificial Intelligence (GenAI) is now used in one form or the other throughout the world and for a variety of purposes. In the process of teaching and learning, the use of GenAI has seen an unprecedented spread in higher education, mainly due to the

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 125–140, 2025.
https://doi.org/10.1007/978-3-031-93539-8_9

large degree of possible adaptation and the considerable number of functionalities that it possesses [1, 2]. As the number of GenAI tools increases at a rapid pace, understanding the perceptions of EAP teachers about them becomes substantial to the process of teaching and learning in higher education [3]. EAP subjects in many universities around the world are taken by students in their first year, therefore, can be used to set trends and standards required as part of institutional quality assurance [4]. The use of GenAI by EAP teachers could be seen as innovative and challenging depending on the personal and academic mindsets of these educators [5].

While educators are commonly viewed and considered to be promoters of the avant-garde, exploring their in-depth perceptions about the use of these tools in EAP classes can provide valuable insights into their effectiveness and ethical implications [4]. The exploration of teachers' perceptions should take into consideration the use of GenAI tools for the main aspects of teaching and learning such as instructional and assessment design that ensure adequate curriculum alignment [6]. Furthermore, it is important to explore the perceptions of educators about the benefit of using these tools to increase students' active engagement and subsequently improve their academic performance.

Often, technological advances are seen as a progressive force that not only creates new frontiers of research but most importantly empowers people with tools that facilitate development and growth [1]. One very important factor that requires appropriate attention by researchers and educators is the need to contextualise the use of these tools appropriately, aiming to purposefully improve students' learning experience and supports their personal and academic growth, while adhering to the notions of academic integrity [3, 7]. Another crucial factor that requires consideration is the acceptance of the benefits of these tools, which ideally should start by getting to know the tools and experimenting with them to ensure faculty familiarisation with their functions [6]. The pedagogical implications of using GenAI tools in EAP classes necessitates their use based on the potential benefits that they offer to teach aspects of vocabulary, grammar, sentence formation, mechanics of academic writing and critical and analytical thinking [4, 7].

Nevertheless, utilisation of GenAI tools in EAP classes brings along a set of concerns and worries that educators may have, which are connected with individual, technological and ethical sets of challenges [7]. Addressing all these challenges may seem a daunting task, nevertheless, exploring what they are and how they make educators feel, serves the purpose of their successful implementation in the process of teaching and learning in higher education. Concerns of educators could be grouped into three main categories: a) scepticism about the benefits of GenAI tools, b) inability to understand and explore these tools efficiently, and c) the ethical implications connected with the use of these tools.

This paper aims to explore the perceptions, use and ethical implications of GenAI in academic writing from the perspectives of EAP faculty with the following research questions proposed:

1. How do EAP faculty perceive the integration of GenAI into their pedagogical practices?
2. What benefits and challenges are perceived by the faculty regarding GenAI usage?
3. How are ethical considerations shaping faculty attitudes towards GenAI?

By investigating these questions, this study aims to provide valuable insights for faculty and higher education management, aiming to positively influence instructional methods, assessment tasks, as well as institutional policies. Additionally, this research contributes to the ongoing discussion on EAP teachers' use of GenAI tools, which could equip faculty with a deeper understanding of potential challenges and gaps in teaching, assessment, and overall students' academic performance. By specifically exploring the perspectives of EAP teachers, this study explores the perceptions of significant stakeholders in higher education and an area with limited empirical research.

2 Literature Review

2.1 Technology Acceptance Model in Higher Education

The Technology Acceptance Model (TAM) is based on two main characteristics; perceived usefulness and perceived ease in using the technology [8, 9]. While the origin of this widely accepted and researched theoretical framework dates back to the 1980s, its spread and especially the diversity of GenAI applications makes it a suitable theoretical framework that can be utilised to explore and analyse the perceptions of faculty about the ease in using the tools and their degree of usefulness [10]. GenAI is perceived to have its benefits and drawbacks when considering its use in the teaching and learning process in higher education [11]. Understandably faculty in higher education consider all the aspects of teaching and learning when explaining their perceptions about the benefits or the drawbacks of technological tools [12, 13]. These aspects are related to instructional, curriculum and assessment design. Although the same tools may not necessarily be utilised in all these three major components of teaching and learning, the usefulness and easiness in the perceived understanding remain equally important to be explored for the benefit of the teaching and learning process.

Exploration of perceived usefulness and ease of EAP faculty in using GenAI tools to facilitate teaching and learning in higher education relates closely to understanding the features of these tools as well as the ways that may be utilised in the process of teaching and learning. While it is understandable that different faculty members may have different approaches when considering the use of GenAI in their classrooms, it is important to explore their perceived understanding of the use of these tools to increase the efficacy of their implementation process [10]. Implementation of GenAI requires not only a certain degree of understanding of these tools, but most importantly requires a holistic acceptance of their unique features that can enrich students' learning experience and subsequently create a life-long learning experience [14]. Exploration of GenAI tools has the power to impact the process of knowledge transference from faculty to students. While accepting the variety of benefits that these tools possess, it is important to recognise the issues that may arise with regard to their ethical use in higher education. Therefore, faculty in EAP classes need to approach the use of GenAI tools with an open mind and understand and analyse any issues related to their use and how this may affect academic integrity.

2.2 EAP Faculty and Technological Adoption

The evolution of technology in EAP reflects a broader shift in educational practices. Early implementations focused on audio-visual tools and the use of language labs, aimed at improving listening and pronunciation skills. By the 1980s and 1990s, computer-assisted language learning (CALL) systems became prevalent, offering interactive exercises and real-time feedback. These tools represented a significant leap, emphasising individualised learning approaches [15]. The advent of the internet in the 2000s marked a turning point, with online learning platforms such as Blackboard allowing EAP educators to incorporate multimedia resources, asynchronous learning, and digital communication tools into their teaching [16]. Mobile learning technologies further enhanced this evolution, enabling students to access resources on the go through apps focused on vocabulary building and grammar enhancement. Advancements centred on AI and machine learning where platforms such as Grammarly provide real-time language corrections and stylistic suggestions, allowing students to refine their work autonomously [17]. However, GenAI tools like ChatGPT and Claude have now gained prominence for their ability to generate coherent, context-specific text, aiding students in drafting essays and improving academic writing skills [18].

The adoption of these advanced tools by EAP faculty is influenced by several factors. The TAM highlights that perceived ease of use and alignment with pedagogical goals are critical determinants [19]. Therefore, faculty are more likely to integrate technologies that offer tangible benefits in terms of student engagement and outcomes. Institutional support, access to resources, professional development opportunities, and a supportive culture encouraging experimentation with new technologies also play a pivotal role in facilitating adoption [20]. However, ethical concerns shape faculty attitudes, including data privacy and algorithmic bias. According to Holmes et al. [21], transparency and ethical guidelines in AI integration are critical to addressing these challenges and building trust among educators and learners.

2.3 Impact of GenAI on EAP Practices

GenAI is now profoundly influencing curriculum design and delivery of EAP, enabling dynamic, personalised, and flexible instructional methods. AI-driven platforms support educators in creating tailored materials for diverse learner needs, offering resources that align with specific academic goals [18]. These tools allow students to practice drafting essays, refine arguments, and receive immediate feedback on grammar and structure, fostering autonomous learning. Educators are also able to analyse student performance data to customise learning paths, ensuring targeted interventions for areas of improvement. These tools help address individual student challenges in real-time, enhancing learning outcomes [22]. GenAI also enhances blended learning approaches, allowing seamless integration of traditional instruction with digital tools. This flexibility is particularly beneficial for diverse classrooms, supporting synchronous and asynchronous learning modes. However, Ansor et al. [23] suggest that the successful application of these tools requires careful alignment with pedagogical objectives and ongoing evaluation to ensure they effectively meet educational goals.

GenAI has the potential to revolutionise assessment and feedback mechanisms in EAP by streamlining processes and providing detailed, consistent evaluations. According to Hu et al. [24], automated tools can assess complex tasks such as essays and presentations with precision, saving time and reducing faculty workload These systems not only evaluate grammar and syntax but also offer feedback on critical thinking and coherence, essential components of academic writing. AI also supports formative assessments through tools that provide real-time feedback to students, assessing oral fluency and accuracy. Such tools encourage iterative learning, allowing students to address weaknesses immediately [17]. Despite its benefits, the integration of AI in assessment raises concerns about data privacy, algorithmic biases, and the inability to assess creative or nuanced language use effectively. Faculty often combine AI-generated evaluations with manual reviews to ensure a holistic and fair assessment [21]. Hence, ethical considerations, such as transparency in data usage and adherence to privacy laws, are critical in addressing these challenges [25].

2.4 Faculty Development and Support

The effective integration of GenAI into educational settings significantly hinges on substantial faculty development and robust institutional support. This aspect extends beyond mere technological training to encompass a holistic approach combining AI tools with pedagogical strategies. Research highlights the significance of targeted professional development programmes that enhance faculty competencies and confidence in using advanced technologies [20, 26–28]. The necessity for faculty development stems from ensuring that educators are technically proficient and pedagogically adept at integrating AI into their teaching methodologies effectively [29, 30]. Professional development initiatives must address the technical/technological skills required to use AI tools and pedagogical strategies to leverage these technologies effectively. Furthermore, professional development programmes must focus on bridging the gap between the potential of AI technologies and the current practices of faculty, emphasising the creation of dynamic, interactive learning environments that enhance teaching and learning experiences.

Institutional support structures are equally crucial in successfully adopting AI in educational settings, as they provide the structural and resource-based backbone necessary for the broad adoption of AI technologies. This support should include access to AI tools, ongoing training programmes, and a supportive culture encouraging innovation and experimentation [31]. Bobrytska et al. [32] emphasise stakeholder engagement as another critical component of institutional support. Their research indicates that the attitudes and expectations of various stakeholders, including students, faculty, and administrators, play a vital role in the acceptance and effectiveness of AI tools in higher education and that institutions that provide comprehensive support in resources, training time, and collaborative learning opportunities see a higher rate of successful AI tool integration. This support is particularly crucial in multicultural settings like the UAE, where the diversity of faculty backgrounds may require more personalised and culturally responsive training approaches. Hence, institutions should foster an environment where faculty feel supported in their efforts to adapt and integrate new technologies, enhancing their teaching effectiveness and responsiveness to student needs.

2.5 Challenges and Concerns

While the adoption of GenAI holds tremendous promise for enhancing educational practices, it also brings forth a plethora of challenges and ethical concerns [31]. One of the foremost concerns is AI′s ethical use, particularly in data privacy and the biases inherent in AI algorithms. Chounta et al. [33] and Holmes et al. [21] discuss the importance of developing AI tools that adhere to ethical standards to protect student privacy and ensure fairness in automated decisions. Chan and Tsi [34] highlight that while AI can augment teaching, it cannot replace the human elements of education, such as mentorship and emotional support. Hong et al. [35], Hong et al. [36] and Singh et al. [37], stress that the technological limitations often related to the reliability and sophistication of AI tools must be addressed. Similarly, Wang et al. [22] assert that the effectiveness of AI tools can be influenced by factors such as cultural context and language, which may hinder their applicability to international students. Moreover, pedagogically, AI must align with educational objectives and learning outcomes, a challenge Ansor et al. [23] note is often overlooked in the rush to adopt new technologies. Therefore, educational institutions must establish clear guidelines and moral standards that govern the use of AI technologies to protect faculty and students [25, 38, 39]. This involves ensuring that AI implementations are transparent and adhere to privacy laws and that there are mechanisms to address potential biases or inaccuracies in AI-generated outcomes.

Furthermore, there are technological and pedagogical limitations that need careful consideration. Technologically, the reliability and sophistication of AI tools can vary greatly, impacting their effectiveness and integration into the curriculum [40]. Pedagogically, the challenge lies in aligning AI with educational goals to ensure that it supports learning outcomes rather than detracting from them [41, 42]. Faculty must be equipped not only with the tools but also with the knowledge to integrate these technologies in ways that are pedagogically sound and beneficial to students.

Addressing these challenges requires ongoing research, development, and training to ensure that AI tools are practical and ethically sound and that there is a commitment to continuous improvement and adaptation. Educational institutions must engage in ongoing research and development to refine AI tools, ensuring they meet the evolving needs of academic settings. Simultaneously, they must provide continuous training and support for faculty to keep pace with these technological advancements, ensuring that the integration of AI into teaching and learning is both effective and ethical. Additionally, institutions must engage in comprehensive studies to assess the impact of AI on learning outcomes and to refine these technologies based on empirical evidence [43]. The dynamic interplay between technological innovation and pedagogical practice highlights the complex landscape of integrating AI in education, underscoring the need for a thoughtful and coordinated approach and the development of frameworks that evaluate the performance of AI in promoting sustainable learning outcomes [44].

3 Methodology

3.1 Theoretical Framework

This research is anchored in the Technology Acceptance Model (TAM), which elucidates how users accept and utilise technology. The TAM is particularly relevant in analysing how EAP faculty at multicultural universities approach the adoption of GenAI, focusing on perceived ease of use and perceived usefulness [45]. This study extends the TAM by integrating factors pertinent to multicultural educational settings, reflecting such environments' unique challenges and dynamics. Further, the framework incorporates additional variables such as cultural adaptability and institutional support, which are critical in understanding technology acceptance in diverse academic landscapes.

3.2 Research Design

This research′s qualitative case study design facilitates a detailed and nuanced examination of EAP faculty perspectives towards GenAI at a multicultural university in the UAE. This approach supports an in-depth exploration of how faculty integrate GenAI into their pedagogical practices within the broader institutional and cultural context.

3.3 Aims and Objectives

This research aims to delve into the perceptions and readiness of EAP faculty to integrate GenAI tools into their teaching. The objectives are manifold:

1. To identify the perceived benefits and barriers to GenAI adoption in EAP settings.
2. To understand the ethical considerations that influence faculty attitudes toward GenAI.
3. To evaluate the impact of GenAI on pedagogical practices and student outcomes.

3.4 Sample Selection

The sample includes 20 EAP faculty members at a multicultural university in the UAE, all of whom have used GenAI tools in their teaching. Purposive sampling was employed to ensure a diverse mix of ages, experiences, and cultural backgrounds, reflecting the rich diversity of the faculty at the university.

3.5 Data Collection Procedures

Data was collected using digital written interviews to enable a comprehensive and reflective exploration of EAP faculty perspectives on GenAI. These interviews were facilitated by distributing Word documents containing open-ended questions to elicit detailed insights into the participants' experiences and perceptions. This format was selected to accommodate the faculty members' diverse schedules and availability, allowing them ample time to provide thoughtful and in-depth responses without the constraints of synchronous communication.

Each participant received the interview document via email, along with detailed instructions on completing and returning their responses. To ensure consistency and

reduce potential confusion, the document included clear guidelines on the length and depth of responses, assurances of confidentiality and instructions for anonymising personal data. Participants were given a two-week window to return their completed documents, providing sufficient time to engage deeply with the questions posed.

3.6 Data Analysis Procedures

Upon collection, the data underwent a rigorous thematic analysis process to uncover patterns and themes related to the adoption and impact of GenAI within EAP pedagogical practices. Initially, the research team thoroughly read all returned documents multiple times to gain a deep understanding of the content, beginning to note initial ideas and patterns. Responses were then coded line by line using qualitative data analysis software. Initial codes reflecting the content were generated and consistently applied across the dataset, with multiple researchers involved to ensure intercoder reliability and mitigate bias. These codes were collated into potential themes and subthemes by grouping similar codes and analysing how they collectively addressed the research questions.

The possible themes were meticulously reviewed and refined, involving a back-and-forth examination between the dataset and coded extracts to fine-tune each theme′s specifics and confirm their accurate representation of the data. Themes were also evaluated against each other to ensure they were distinct yet formed a coherent narrative. After finalising the themes, they were clearly defined and named to reflect the data′s depth and complexity accurately. The final analysis stage involved selecting vivid and compelling extract examples that best illustrated each theme integrating these into the analytical narrative. The comprehensive report contextualised these findings within the adapted TAM, explicitly tailored to the multicultural educational context of this study. The research team maintained a rigorous audit trail throughout the analysis to document all decisions made during the coding and theme development processes. This trail includes notes on why specific themes were merged or discarded, reflections on the coding process, and discussions about theme definitions, providing transparency and enhancing the trustworthiness of the analysis.

This comprehensive methodological framework was designed to offer deep insights into the complex dynamics of GenAI adoption among EAP faculty in multicultural educational settings. The structured approach aimed to contribute significantly to the discourse on integrating innovative technologies in higher education, particularly in environments characterised by diverse academic needs and cultural backgrounds. Moreover, the detailed data collection and analysis approach ensured that the study captured a nuanced understanding of how EAP faculty perceive and integrate GenAI into their teaching, focusing mainly on the ethical, practical, and pedagogical dimensions of technology use in a multicultural academic environment.

4 Findings and Analysis

The findings of this study are structured to address the above research questions. By investigating these questions, this study aims to provide actionable insights for educators, policymakers, and institutions striving to balance innovation with ethical and pedagogical integrity in the age of AI.

4.1 Faculty Perceptions of GenAI in Pedagogical Practices

EAP faculty reported varied perceptions of GenAI tools, reflecting differences in their familiarity with and confidence in the technology. GenAI tools such as ChatGPT, Gemini, and Copilot were employed for a range of tasks, including generating teaching materials, simplifying complex texts, and automating routine administrative functions. For some instructors, these tools represented a revolutionary shift in pedagogical practices. One respondent noted, *"AI has streamlined my teaching process, making it possible to focus on engaging with students rather than spending hours on administrative work."*

Many instructors highlighted the potential of GenAI to support differentiated instruction, enabling them to address the diverse needs of students in a multicultural academic context. For instance, one participant explained, *"AI allows me to create multiple versions of the same task, ensuring that all students, regardless of their proficiency level, have access to appropriate learning materials."* This adaptability was particularly valued in classrooms with mixed-ability learners, where instructors often struggled to cater to individual needs within tight time constraints.

However, faculty attitudes toward GenAI were not universally positive. While some embraced its capabilities with optimism, others adopted a more cautious stance, using the tools selectively and ensuring human oversight at every stage. One instructor remarked, *"AI is a useful assistant, but it's not infallible. I always review the materials it generates to ensure they meet the required standards."* This cautious acceptance reflected a recognition of GenAI's potential, tempered by concerns about its limitations. A smaller group of faculty members expressed scepticism or resistance to GenAI, citing concerns about its reliability, ethical implications, and the potential erosion of fundamental teaching skills. One sceptical participant stated, *"AI may automate tasks, but it cannot replicate the judgment and expertise of a teacher. Over-reliance on technology risks diluting the essence of education."*

4.2 Benefits and Challenges of GenAI Usage

The benefits of GenAI tools were widely acknowledged, with faculty identifying improvements in efficiency, personalisation, and student engagement as key advantages. GenAI significantly reduced the time required for routine teaching tasks, enabling instructors to focus on higher-order pedagogical activities. Faculty reported that tasks such as lesson planning, assessment design, and generating feedback could be completed in a fraction of the time previously required. One instructor commented, *"With AI, I can create a full set of quiz questions in minutes. This efficiency has allowed me to devote more time to mentoring students and refining my teaching strategies."* Efficiency gains were particularly valued in resource-constrained environments, where faculty often manage large classes with diverse needs. The automation of repetitive tasks was seen as a means to alleviate administrative burdens, freeing instructors to focus on student-centred teaching practices.

Another widely reported benefit was the ability of GenAI tools to support personalised learning. Faculty used these tools to tailor tasks and materials to suit individual student needs, creating differentiated learning experiences. For example, one instructor explained, *"AI enables me to generate activities that cater to different proficiency*

levels, ensuring that all students are appropriately challenged." This capability was especially beneficial in multicultural classrooms, where linguistic and cultural diversity often necessitated customised teaching approaches.

GenAI was also credited with increasing student engagement, particularly among tech-savvy learners who viewed the integration of AI into their education as innovative and exciting. One respondent remarked, *"Students are genuinely excited to explore how AI can support their learning. This curiosity has translated into more active participation in discussions and activities."* Faculty noted that GenAI's ability to generate dynamic and interactive content made lessons more engaging, fostering a sense of relevance and modernity in the classroom. For example, tools like ChatGPT were used to create simulated academic conversations, encouraging students to experiment with language and ideas in a low-pressure environment.

Despite these benefits, the integration of GenAI tools was not without challenges. Faculty identified several obstacles, ranging from technical limitations to broader ethical and equity concerns. A key concern was the variable quality of AI-generated outputs. Faculty noted that while these tools could produce content quickly, the outputs often required careful review and adaptation to ensure they met academic standards. One instructor shared, *"AI-generated materials can look polished, but they often lack the depth and critical perspective needed for higher-level academic tasks."* This limitation necessitated additional time and effort on the part of instructors, undermining some of the efficiency gains provided by GenAI.

Many instructors expressed concerns about the potential for students to become overly dependent on GenAI tools, bypassing critical thinking and independent problem-solving processes. One respondent observed, *"Students find it tempting to rely on AI for everything, but this approach doesn't help them develop the skills they need for success in the long term."* Faculty worried that unchecked reliance on AI could erode essential academic competencies, such as analytical reasoning and creativity.

Disparities in access to GenAI tools were highlighted as a significant issue, particularly in diverse academic contexts where not all students have the resources or digital literacy to use these technologies effectively. One participant noted, *"The digital divide is becoming more evident with the rise of AI. Not all students have equal opportunities to benefit from these tools, which creates a gap in learning outcomes."* Faculty also identified a range of ethical and practical challenges, including concerns about plagiarism, academic integrity, and data privacy. One instructor explained, *"AI makes it easier for students to produce polished assignments, but it's also harder to detect plagiarism. This creates a grey area in assessing their work."* Additionally, concerns about the use of sensitive data on AI platforms were frequently mentioned, with faculty calling for clearer institutional guidelines on data security and ethical AI usage.

4.3 Ethical Considerations Shaping Faculty Attitudes Towards GenAI

Ethical considerations played a central role in shaping faculty attitudes toward GenAI, influencing how they approached its integration into their teaching practices. Faculty expressed significant concerns about the impact of GenAI on academic integrity. Instances of students using AI tools to generate essays and assignments were frequently reported, raising questions about how to uphold originality and effort in an era where

AI is readily accessible. One instructor stated, *"It's becoming increasingly difficult to distinguish between authentic student work and AI-generated content. This challenges our ability to assess learning accurately."*

Another major ethical concern was the potential misuse of sensitive student data on AI platforms. Faculty were cautious about uploading institutional or student information, citing the need for secure and compliant platforms. One respondent remarked, *"Before we integrate AI fully, we need to address the privacy implications and ensure that the tools we use are ethically sound."* Equity issues were also prominent, with faculty noting that students from less privileged backgrounds often lacked access to GenAI tools or the skills to use them effectively. This disparity created challenges in fostering inclusive and equitable learning environments. As one participant explained, "*AI should be a tool that enhances education for everyone, but without proper support, it risks becoming a barrier for some students."*

4.4 Strategies for Addressing Challenges

Faculty adopted various strategies to mitigate the challenges associated with GenAI. Transparency in AI usage was a key approach, with instructors encouraging students to disclose when they had used AI tools in their assignments. Additionally, many faculty redesigned assessments to focus on process-oriented tasks, such as multi-stage assignments, in-class activities, and oral presentations, to ensure authenticity and reduce the risk of plagiarism.

Another strategy was the promotion of digital literacy among students. Faculty introduced workshops to educate learners about the strengths and limitations of GenAI, as well as its ethical implications. One instructor noted, *"Helping students understand when and how to use AI responsibly is just as important as teaching them the course content."* Institutional support was identified as critical for addressing broader issues of equity and ethics. Faculty called for clearer policies on AI usage, access to secure tools, and resources to support both instructors and students in navigating this new technological landscape.

5 Discussion

This study highlighted the complexity of utilising GenAI tools in EAP classes in higher education, especially relating to pedagogical, practical, and ethical implications, and a direct relationship to faculty preparedness to alleviate this tension. Grounded in the Technology Acceptance Model (TAM), the analysis shows that faculty perceptions of GenAI are significantly influenced by their perceived usefulness and ease of use, which are central tenets of the model. Faculty perceptions throughout the study were situated within three interconnected themes: Optimistic Integration (OI), Cautious Acceptance (CA), and Resistance and Scepticism (RS), reflecting diverse attitudes toward GenAI adoption. By directly applying the TAM framework, this discussion underscores the importance of addressing the technological and pedagogical aspects of GenAI use, ensuring that such technologies are practical, user-friendly, and aligned with educational goals.

Most of the faculty exhibited OI, viewing GenAI as beneficial due to its perceived high utility in enhancing engagement, personalising learning, and developing EAP skill sets. This aligns with the TAM′s concept of perceived usefulness, where faculty believe that GenAI tools will improve their teaching effectiveness and student outcomes. These findings agree with the results presented by Roe et al. [46] and Wang and Dang [47], which highlight the need to consider faculty perceptions when adopting GenAI in higher education. Studies by Mezei [48] and Roe and Perkins [49] similarly underscored the importance of personalisation in education, echoed here in the value of GenAI tools for tailoring the learning experience. However, it was also noted that while active engagement with GenAI tools is not always translated into efficacy in learning, there is evidence of positive impacts on students' attention span and their understanding of topics, which is similar to the findings by Bannister [6].

Conversely, the theme of CA illustrates a critical engagement with the perceived usefulness and ease of use of GenAI. Those in this category acknowledge the potential benefits of GenAI but are also acutely aware of the practical hurdles that could impede its effective integration. This reflects the TAM′s assertion that acceptance is likely when users acknowledge the benefits of a technology but also consider its usability and integration challenges. Faculty remain mindful of inevitable challenges such as accuracy, over-reliance, equity, and, most importantly, ethics. Faculty were eager to see institutional policy changes and the inclusion of GenAI in these policies to ensure quality assurance, involving students, faculty, and policymakers. This differs slightly from the findings by Roe et al. [46], who argue that policies need to guide GenAI adaptations in higher education, particularly when considering assessment design and ethical implications. Tantivejakul et al. [50] also view the adoption of GenAI in EAP as a complex process that requires alignment with institutional and disciplinary policies, which this study affirms.

Lastly, the RS theme arose from concerns over the ease of use and doubts about the actual utility of GenAI, which are directly linked to the TAM′s factors. Sceptical faculty members question whether the technology can genuinely enhance educational outcomes or complicate the teaching process. In some cases, faculty address the risks related to ethical concerns and possible practical, linguistic, or cultural barriers. These perspectives articulate similar faculty perceptions as those highlighted by Bannister and Carver [51], although they disagree to a certain extent with the findings by Roe et al. [46], who perceive faculty concerns as being primarily due to a lack of understanding of the importance of GenAI rather than the difficulties in adapting it. Kohnke [52] adds to this by highlighting that a lack of practical knowledge, such as constructing effective prompts, can contribute to such scepticism.

While this study found that students' engagement levels and self-efficacy increase due to the use of GenAI tools, it also realised that in the process of adapting these tools, it is important to focus on the practical challenges faced by students who are not equipped with the knowledge of using the tools, are less technologically advanced than their peers, or lack practical knowledge such as the ability to construct effective prompts [52]. These challenges also resonate across the spectrum of CA and RS, as they require faculty to consider strategies to bridge the gap. This study agreed that to ensure efficacy in the use of GenAI tools, strategies addressing challenges faced by EAP faculty, transparency,

redesigned assessments, and digital literacy initiatives are required to be implemented to ensure effective and successful integration of GenAI in diverse educational contexts, which is similar to the findings by Zhang and Dong [53]. Finally, the study echoed other research highlighting that GenAI use in EAP classes can increase faculty's understanding of the ethical and effective use of GenAI for the benefit of learning and not the opposite.

6 Conclusion

This study has systematically explored the perceptions, experiences, and challenges faced by EAP faculty at a multicultural university in the UAE regarding integrating GenAI into their teaching practices. Grounded in the Technology Acceptance Model (TAM), the study highlights how perceived usefulness and ease of use significantly influence faculty attitudes towards GenAI adoption. The qualitative analysis of digital written interviews revealed a spectrum of responses ranging from Optimistic Integration to Cautious Acceptance and Resistance and Scepticism, each reflecting diverse attitudes towards GenAI′s role in education. While there is significant optimism about the potential of GenAI to transform educational delivery by enhancing efficiency, personalising learning experiences, and fostering student engagement, there are also profound concerns. These concerns primarily revolve around the ethical implications, the fear of depersonalisation of education, potential job losses, and the erosion of critical humanistic values. Moreover, this study has highlighted the need for comprehensive faculty development programmes that focus on the technical aspects of GenAI and address pedagogical strategies to ensure that technology enhances teaching and learning without compromising educational values. Such programmes should be sensitive to the multicultural setting of the institution, recognising the diverse cultural backgrounds of faculty and students.

In conclusion, while GenAI offers substantial opportunities for advancing educational practices, its integration into the curriculum must be handled carefully, ensuring alignment with pedagogical goals and ethical standards. Institutions must adopt a balanced approach that promotes technological advancements while safeguarding the core values and integrity of the educational process. This study contributes to the broader discourse on the use of AI in education, offering insights that could guide policymakers, curriculum developers, and educational technology designers in creating environments that leverage GenAI′s potential responsibly and effectively.

References

1. Guillén-Yparrea, N., Hernández-Rodríguez, F.: Unveiling generative AI in higher education: insights from engineering students and professors. In: 2024 IEEE Global Engineering Education Conference (EDUCON), pp. 1–5. IEEE (2024)
2. Wang, H., Dang, A., Wu, Z., Mac, S.: Generative AI in higher education: seeing ChatGPT through universities' policies, resources, and guidelines. Comput. Educ. Artif. Intell. **7**, 100326 (2024)
3. Hoernig, S., Ilharco, A., Pereira, P.T., Pereira, R.: Generative AI and Higher education: challenges and opportunities (2024)

4. Ngo, T.N., Hastie, D.: Artificial intelligence for academic purposes (AIAP): integrating AI literacy into an EAP module. Engl. Specif. Purp. **77**, 20–38 (2025)
5. Fan, Y., Tan, S., Lim, G.Y.W.: EAP teacher feedback in the age of AI: supporting first-year students in EFL disciplinary writing. Aust. J. Appl. Linguist. **7**(3), 1943 (2024)
6. Bannister, P.: English medium instruction educator language assessment literacy and the test of generative AI in online higher education (2024)
7. Şık Keser, K., Razı, S.: Upholding academic integrity in EAP. In: Teaching English for Academic Purposes: Theory into Practice, pp. 321–342. Cham: Springer Nature Switzerland (2024)
8. Marangunić, N., Granić, A.: Technology acceptance model: a literature review from 1986 to 2013. Univ. Access Inf. Soc. **14**, 81–95 (2015). https://doi.org/10.1007/s10209-014-0348-1
9. Silva, P.: Davis' technology acceptance model (TAM) (1989). Inf. Seeking Behav. Technol. Adoption: Theor. Trends, 205–219 (2015)
10. Davis, F.D., Granić, A.: The technology acceptance model: 30 years of TAM. Springer International Publishing AG (2024). https://doi.org/10.1007/978-3-030-45274-2
11. Kong, S.C., Yang, Y., Hou, C.: Examining teachers' behavioral intention of using generative artificial intelligence tools for teaching and learning based on the extended technology acceptance model. Comput. Educ. Artif. Intell. **7**, 100328 (2024). https://doi.org/10.1016/j.caeai.2024.100328
12. Davis, F.D.: User acceptance of information systems: the technology acceptance model (TAM) (1987)
13. Lala, G.: The emergence and development of the technology acceptance model (TAM). Mark. Inf. Decis. **7**, 149–160 (2014)
14. Ghimire, A., Edwards, J.: Generative AI adoption in classroom in context of technology acceptance model (TAM) and the innovation diffusion theory (IDT). arXiv preprint arXiv:2406.15360. https://doi.org/10.48550/arXiv.2406.15360
15. Levy, M.: Computer-assisted language learning: context and conceptualization (2023). https://doi.org/10.1093/oso/9780198236320.001.0001
16. Stockwell, G., Wang, Y.: Exploring the challenges of technology in language teaching in the aftermath of the pandemic. RELC J. **54**(2), 474–482 (2023). https://doi.org/10.1177/00336882231168438
17. Farr, F., Murray, L.: The routledge handbook of language learning and technology (1st ed.). Routledge (2016). https://doi.org/10.4324/9781315657899
18. Li, Y.: Study on the role and influence of AI in personalized learning for second language acquisition. J. Educ. Humanit. Soc. Sci. **38**, 125–131 (2024). https://doi.org/10.54097/wfe3gk31
19. Davis, F.D., Venkatesh, V.: Technology acceptance in education revisited: implications for faculty adoption. J. Inf. Syst. Educ. **34**(1), 1–15 (2023)
20. Keese, J., Ford, D., Luke, S., Vaughn, S.: An individualized professional development approach for training university faculty in using technological tools. Educ. Inf. Technol. **28**(11), 14577–14594 (2023). https://doi.org/10.1007/s10639-023-11792-8
21. Holmes, W., et al.: Ethics of AI in education: towards a community-wide framework. Int. J. Artif. Intell. Educ. **32**, 504–526 (2022). https://doi.org/10.1007/s40593-021-00239-1
22. Wang, T., et al.: Exploring the potential impact of artificial intelligence (AI) on international students in higher education: generative AI, chatbots, analytics, and international student success. Appl. Sci. **13**(11), 6716 (2023). https://doi.org/10.3390/app13116716
23. Ansor, F., Zulkifli, N.A., Jannah, D.S.M., Krisnaresanti, A.: Adaptive learning based on artificial intelligence to overcome student academic inequalities. J. Soc. Sci. Utilizing Technol. **1**(4), 202–213 (2023). https://doi.org/10.55849/jssut.v1i4.663

24. Hu, S., Ke, F., Vyortkina, D., Hu, P., Luby, S., O'Shea, J.: Artificial intelligence in higher education: applications, challenges, and policy development and further considerations. In: Perna, L.W. (eds.) Higher Education: Handbook of Theory and Research. Higher Education: Handbook of Theory and Research, vol. 40. Springer (2025). https://doi.org/10.1007/978-3-031-51930-7_13-2
25. Selvaratnam, R., Venaruzzo, L.: Governance of artificial intelligence and data in higher education: a snapshot of policy and practice. ASCILITE Publ. 1–7 (2024). https://doi.org/10.14742/apubs.2023.717
26. Ibrahim, A.B.: Assessing the knowledge and perception of artificial intelligence for teaching and research among lecturers in the faculties of arts in Nigeria. J. Glob. Res. Educ. Soc. Sci. **18**(2), 25–33 (2024). https://doi.org/10.56557/jogress/2024/v18i28671
27. Lee, J., Wu, A.S., Li, D., Kulasegaram, K.: Artificial intelligence in undergraduate medical education: a scoping review. Acad. Med. **96**(11S), S62–S70 (2021). https://doi.org/10.1097/acm.0000000000004291
28. Russell, R.G., et al.: Competencies for the use of artificial intelligence–based tools by health care professionals. Acad. Med. **98**(3), 348–356 (2022). https://doi.org/10.1097/acm.0000000000004963
29. Arvin, N., Hoseinabady, M., Bayat, B., Zahmatkesh, E.: Teacher experiences with AI-based educational tools. AI Tech in Behav. Soc. Sci. **1**(2), 26–32 (2023). https://doi.org/10.61838/kman.aitech.1.2.5
30. Nguyen, G., Nguyen, N.P., Giang, N.T.H.: Situation and proposals for implementing artificial intelligence-based instructional technology in Vietnamese secondary schools. Int. J. Emerg. Technol. Lear. (iJET) **17**(18), 53–75 (2022). https://doi.org/10.3991/ijet.v17i18.31503
31. Slimi, Z., Villarejo-Carballido, B.: Unveiling the potential: experts perspectives on artificial intelligence integration in higher education. Eur. J. Educ. Res. **13**(4), 1477–1492 (2024). https://doi.org/10.12973/eu-jer.13.4.1477
32. Bobrytska, V.I., Krasylnykova, H.V., Beseda, N.A., Krasylnykov, S.R., Skyrda, T.S.: Artificial intelligence (AI) in Ukrainian higher education: a comprehensive study of stakeholder attitudes, expectations and concerns. Int. J. Learn., Teach. Educ. Res. **23**(1), 400–426 (2024). https://doi.org/10.26803/ijlter.23.1.20
33. Chounta, I., Bardone, E., Raudsep, A., Pedaste, M.: Exploring teachers' perceptions of artificial intelligence as a tool to support their practice in Estonian K–12 education. Int. J. Artif. Intell. Educ. **32**(3), 725–755 (2021). https://doi.org/10.1007/s40593-021-00243-5
34. Chan, C.K.Y, Tsi, L.H.Y.: The AI revolution in education: will AI replace or assist teachers in higher education? ArXiv, abs/2305.01185 (2023). https://doi.org/10.48550/arXiv.2305.01185
35. Hong, Y., et al.: Statistical perspectives on reliability of artificial intelligence systems (2021). https://doi.org/10.48550/arxiv.2111.05391
36. Hong, Y., Min, J., King, C., Meeker, W.Q.: Reliability analysis of artificial intelligence systems using recurrent events data from autonomous vehicles. J. R. Stat. Soc. Ser. C. Appl. Stat. **71**(4), 987–1013 (2022). https://doi.org/10.1111/rssc.12564
37. Singh, R., Vatsa, M., Ratha, N.K.: Trustworthy AI. Proceedings of the 3rd ACM India Joint International Conference on Data Science & Management of Data (8th ACM IKDD CODS), pp. 449–453 (2021). https://doi.org/10.1145/3430984.3431966
38. Barnes, E., Hutson, J.: Navigating the ethical terrain of AI in higher education: Strategies for mitigating bias and promoting fairness. Forum Educ. Stud. **2**(2), 1229 (2024). https://doi.org/10.59400/fes.v2i2.1229
39. Ma, X., Jiang, C.: On the ethical risks of artificial intelligence applications in education and its avoidance strategies. J. Educ. Humanit. Soc. Sci. 14, 354–359 (2023). https://doi.org/10.54097/ehss.v14i.8868

40. Barua, P.D., et al.: Artificial intelligence enabled personalised assistive tools to enhance education of children with neurodevelopmental disorders—a review. Int. J. Environ. Res. Public Health **19**(3), 1192 (2022). https://doi.org/10.3390/ijerph19031192
41. Opesemowo, O.A.G., Adekomaya, V.: Harnessing artificial intelligence for advancing sustainable development goals in South Africa′s higher education system: a qualitative study. Int. J. Learn. Teach. Educ. Res. **23**(3), 67–86 (2024). https://doi.org/10.26803/ijlter.23.3.4
42. Veras, M., Dyer, J., Rooney, M., Barros Silva, P.G., Rutherford, D., Kairy, D.: Usability and efficacy of artificial intelligence chatbots (ChatGPT) for health sciences students: protocol for a crossover randomized controlled trial. JMIR Res. Protoc. **12**, e51873 (2023). https://doi.org/10.2196/51873
43. Gligorea, I., Cioca, M., Oancea, R., Gorski, A., Gorski, H., Tudorache, P.: Adaptive learning using artificial intelligence in e-learning: a literature review. Educ. Sci. **13**(12), 1216 (2023). https://doi.org/10.3390/educsci13121216
44. Yman, C., Mohamed, A.: Fostering sustainable development through artificial intelligence in education (ESD): a comprehensive evaluation framework and key criteria analysis (2023). https://doi.org/10.21203/rs.3.rs-3372485/v1
45. Kanont, K., et al.: Generative-AI, a learning assistant? factors influencing Higher-Ed students technology acceptance. Electron. J. e-Learning **2**(6), 18–33 (2024). https://doi.org/10.34190/ejel.22.6.3196
46. Roe, J., Perkins, M., Tregubova, Y.: The EAP-AIAS: adapting the AI assessment scale for english for academic purposes. arXiv preprint arXiv:2408.01075 (2024). https://doi.org/10.48550/arXiv.2408.01075
47. Wang, H., Dang, A.: Enhancing L2 writing with generative AI: a systematic review of pedagogical integration and outcomes (2024). https://doi.org/10.13140/RG.2.2.19572.16005
48. Mezei, A.M.: Technology-assisted teaching activities in an ESAP context: an ecological perspective. Stud. Univ. Babes-Bolyai-Philologia **69**(4), 17–43 (2024)
49. Roe, J., Perkins, M.: Generative AI in self-directed learning: a scoping review. arXiv preprint arXiv:2411.07677 (2024). https://doi.org/10.48550/arXiv.2411.07677
50. Tantivejakul, N., Chantharasombat, J., Kongpolphrom, W.: Voices of the future: exploring students' views on the use of GenAI in academic and professional PR writing. LEARN J. Lang. Educ. Acquisition Res. Netw. **17**(2), 511–537 (2024)
51. Bannister, P., Carver, M.: 'I don't need professional development; I want institutional development': legitimising marginalised epistemic capital that disrupts generative AI discourse. Prof. Dev. Educ. 1–19 (2024)
52. Kohnke, L.: Exploring EAP students' perceptions of GenAI and traditional grammar-checking tools for language learning. Comput. Educ. Artif. Intell. **7**, 100279 (2024). https://doi.org/10.1016/j.caeai.2024.100279
53. Zhang, Y., Dong, C.: Unveiling the dynamic mechanisms of generative AI in English language learning: a hybrid study based on fsQCA and system dynamics. Behav. Sci.. Sci. **14**(11), 1015 (2024). https://doi.org/10.3390/bs14111015

The Challenges Faced by Albanian Teachers in the Use of Media Technology During Teaching

Deniz Çupi and Rudina Alimerko(✉)

Faculty of Humanities, "Ismail Qemali" University, Vlorë, Albania
{deniz.cupi,rudina.alimerko}@univlora.edu.al

Abstract. Media technology is increasingly used in education to transform the teaching and learning process. Tools such as videos, quiz and competition applications, animations, and interactive presentations make teaching more likeable and increase students' engagement. These technologies help them acquire new knowledges faster and more deeply, and contribute to a better understanding of complex concepts. However, the successful use of media technology in the classroom depends on the preparation of teachers to recognize and use these tools. Teachers need ongoing training to help them adapt to the rapid developments in technology and to use these tools effectively. This requires well-thought-out policies, strategies, and investments in the use of media technology in the classroom. This study aims to identify the challenges and needs that Albanian teachers face when using media technology in schools. A survey of teachers will help identify their difficulties and suggest solutions for more effective use. In addition, classroom observations will be conducted where media technology is part of the teaching process. The study is based on several educational theories such as Richard Mayer's multimedia learning theory, Lev Vygotsky's constructivist learning theory, David Kolb's experiential learning theory, Albert Bandura's social learning theory, John Sweller's cognitive load theory, and of many others. Theories and surveys in this field show that multimedia learning is not simply the presentation of information through technology, but a learning system, different from the traditional one. The use of technology in the classroom requires the development of well-defined educational policies and strategies, as emphasized in UNESCO and OECD reports. Ongoings teacher training in the proper use of these technologies plays a critical role in the success of the technology-based instructional process.

Keywords: media technology · teaching · multimedia learning · teachers · training

1 Introduction

The use of media technology in teaching has been an important issue in the field of education in the last 3 decades, and continues to be a key aspect for improving the learning process and engaging students. Technology has enabled significant innovations, enriching learning with new and more efficient opportunities, such as online resources, visual means, and personalized learning platforms. However, the integration of this technology

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 141–154, 2025.
https://doi.org/10.1007/978-3-031-93539-8_10

has encountered various challenges, including the lack of proper infrastructure, insufficient teacher training, and the gaps that exist between public and private schools, as well as between urban and rural ones. In transition countries such as Albania, these challenges are more visible [39].

This study aims to explore the use of media technology in teaching in Albanian schools, evaluating its impact on teaching, the challenges faced by teachers and students, as well as opportunities for improvement and further development. The study focuses on the approach and use of technology by both teachers and students, identifying the benefits and shortcomings related to its integration in the school environment. Also, such analysis will help to understand the differentials that exist between different schools and areas and will provide recommendations for improving the conditions for the efficient use of technology in education. Through this study, it is intended to contribute to the development of educational policies and strategies that enable a sustainable and comprehensive use of technology, creating equal opportunities for all students and teachers across Albania.

The OECD report [26] highlights that the effects of using digital tools for leisure on students' creative thinking differ significantly. In Albania, 33% of students engage in digital leisure activities for over an hour each day during school, which is slightly below the OECD average of 35%. Generally, this usage tends to correlate with a decline in creative thinking performance among students in OECD countries. Conversely, Albanian students who spend over one hour on digital leisure activities outside of school, such as on weekends, achieved scores that were 2.5 points higher than their peers, taking into account gender and the socio-economic profiles of students and schools.

Currently, the development of digital skills of individuals in the Western Balkans is at a low level compared to individuals in western European countries. They possess basic digital skills at least 20% lower than in the EU. "However, while all economies fall below the EU average, there are significant differences: Montenegro is the best performer in the region, with only 7% away from EU levels, while Albania and Kosovo reported rates less than 30% - essentially half of EU levels," the report states [26]. Two years later, the same OECD report [10] estimates that the average number of individuals who have basic or above-basic digital skills in the countries of the European Union is 53.9% of them. In Montenegro this is 47.2% of individuals, in Serbia 41.3%, while in Albania this applies to 23.8% of individuals.

The problem starts with meeting elementary needs, equipping schools with computers. The Higher Education Strategy 2021–2026, [1] it is stated that in 2017, in pre-university schools, the computer-student ratio was 1:27, which is far from the standards of EU countries where 1 computer is available for 3–7 students. On the other hand, the number of non-functional computers reaches 25% of their total number, indicating a lack of maintenance, while the Internet speed, most of the times, does not meet the demands of the users. According to this strategy, the possibilities to use the equipment within the classrooms are very limited, as there are no projectors, wireless network and other tools dedicated to this purpose. All these factors constitute the main obstacle to the use of digital and media technology in our schools. Another obstacle to its use is the lack of coordination at the national level. *"In Albania, there is no central authority that deals with the provision of ICT and Internet services for schools"*, the strategy states. Other

research studies [20, 31] highlight the need for Albania to focus on digitizing education and incorporating innovative technologies. This includes developing digital educational resources, enhancing the professional training of current teachers, and ensuring that future teachers are trained with digital skills [9] Achieving the objectives outlined in the National Strategies and increasing investments by the Albanian government in the education sector, along with related fields, will establish a foundation for advancing smart education in Albania [38].

2 Literature Review

The use of media technology in teaching is supported by several pedagogical and psychological theories that emphasize its effectiveness in involving students and improving learning outcomes. These theories provide a basis for understanding how technology can optimize teaching and learning processes. Multimedia Learning Theory, developed by Richard Mayer, states that learning is most effective when content is presented through a variety of channels, such as text, images, videos, and animations. This approach exploits the brain's ability to process visual and verbal information separately, creating stronger cognitive connections. There is a parallel between McLuhan's formula: Media is the message itself and in our case: Multimedia learning is not just the presentation of information through technology but a learning system, different from the traditional one. Meanwhile, other studies on this issue show [3, 30] that Albania belongs to a typical behaviorist school system and the implementation of modern approaches of constructivism is still facing challenges.

Active Learning Theory states that students learn more effectively when they actively participate in the learning process. Interactive media technologies, such as educational apps, simulators and interactive presentations, engage learners by transforming them from passive recipients to active participants. Lev Vygotsk's Constructive Learning Theory emphasises that knowledge is constructed through social experiences and interactions. Media technology provides dynamic platforms for learners to interact with content and collaborate with peers. Tools such as collaborative software and virtual learning environments support group activities, sharing of ideas and discussions. David Kolb's experiential learning theory suggests that learning through experience, supported by technology, encourages reflection, critical thinking and deeper understanding. Albert Bandura's Social Learning Theory emphasizes the role of observation and imitation in learning. Media technology, especially through videos and animations, provides opportunities for students to observe patterns of behavior, techniques, or problem-solving strategies. Collaborative platforms also foster social interaction, allowing students to learn from peers and share ideas. This strengthens the acquisition of knowledge through shared experiences. John Sweller's Cognitive Load Theory explains that learning is most effective when cognitive overload is minimized. Interactive presentations and multimedia tools simplify complex concepts, improving information retention and understanding by reducing unnecessary cognitive load [2].

2.1 The Issues Addressed

The use of media technology in teaching requires well-structured policies from the Ministry of Education, as well as clear strategies for its implementation. To achieve this goal, sustainable investments in media technology and continuous professional development of teachers through specialized training in this field are needed. To explore this hypothesis, the following research questions are posed: What is the impact of the use of media technology in improving the quality of teaching and learning? What policies and strategies are currently taking place by the Ministry of Education for the integration of media technology in teaching? What are the main challenges faced by schools and teachers in using media technology in the teaching process? How effective are teacher training programs for the use of media technology? What are the needs and expectations of teachers and students regarding the use of media technology in teaching? What kind of investments in technology and infrastructure are needed to support the use of media in pre-university education? How can the use of media technology affect the development of students' digital competences? What are the best examples of media technology use practices in international education and how can these be adapted to the Albanian context? The answers to these questions will lead us to a clear picture of the situation recommending the best solutions.

3 Methodology

This study adopts a mixed methods approach to explore the role of media technology in increasing student engagement and improving learning outcomes. The methodology integrates the collection and analysis of qualitative and quantitative data to ensure a thorough understanding of the research problem.

The study uses a descriptive and exploratory design. Descriptive design is used to determine the extent to which media technology affects student engagement and learning outcomes. Exploratory design is used to explore the perceptions, experiences and challenges of teachers and students regarding media technology in education. The study includes two main groups of participants. A questionnaire was made with 100 teachers in the surrounding towns and villages of the region of Tirana and Vlora at different educational levels (primary, secondary and tertiary) who have incorporated media technology in their teaching practices. Additionally, a questionnaire was done with 100 students from the same institutions and demographic areas to provide their experience in using media technology during the learning process.

Semi-structured interviews were conducted with teachers to gather in-depth insights into their experiences, challenges, and recommendations for improving media technology integration. Discussions were also held with students to explore their perspectives on engagement and learning through media technology.

4 Results

The survey was conducted with 100 teachers from all three levels of education and from different schools, 30 of which were from private schools. The age distribution of the teachers was as follows: 40.6% were aged 31–40, 30.9% were aged 41–50, and

28.5% were aged 51–60. In terms of gender, 57.5% of the respondents were female and 42.5% were male. In addition, 54.6% of the teachers worked in urban schools, while 45.4% came from rural schools. Another questionnaire was conducted with 100 students between the ages of 10 and 18, where 54% were female and 46% were male. Meanwhile, 62% of them attended schools in urban areas and 36% in rural areas, where the number of residents, and therefore of students, is lower.

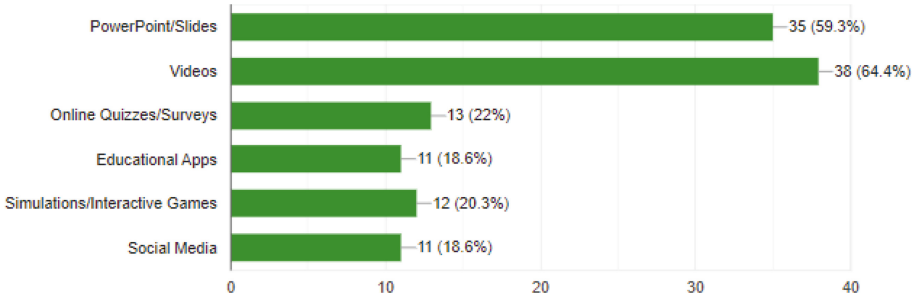
4.1 Teachers and Use of Media Technology in Teaching

According to the survey made by teachers when asked about the technological equipment in their schools, 61.4% reported access to computers/laptops, making this the most common technology; 22.2% had access to tablets/iPads, showing a lower presence of these devices; 40.6% used interactive whiteboards, which are often used for teaching and presentations; 64.4% had projectors, a standard tool in most schools, while 52.6% had access to the internet, a key resource for online materials. Only 13.9% of them used educational apps, highlighting the need for greater investment in this area. The survey asked teachers to evaluate the quality of teaching technology in their schools. Responses revealed: 10.3% rated it as "Excellent," 45.5% as "Good," 25.7% as "Acceptable," and 23.6% as "Poor." While most found the technology adequate (55.8%), a significant portion (49.3%) viewed it as average or poor, highlighting the need for improvements to enhance its quality and effectiveness.

It appears that videos and PowerPoint presentations are the two most commonly used technological tools in teaching, showing a preference for traditional visual and interactive methods. Newer options such as educational apps, interactive games and social networks have yet to be fully embraced by the majority of teachers. Videos are used by 64.4% of teachers, Table 1 making them the most widely used tool. This high usage suggests a strong preference for visual materials to explain concepts and capture student attention. PowerPoint is used by 59.3% of teachers, making it a common choice for visual presentations, helping to organize lesson content in a structured and engaging way. Online quizzes are used by 22% of teachers, reflecting a lower level of adoption but still serving as an essential tool for knowledge assessment and interactive student engagement. Educational apps are used by 18.6%, indicating a more limited use of new technologies in teaching, although they remain present. Simulations and interactive games are used by 20.3% of teachers, showing a relatively low adoption of more advanced tools aimed at developing practical and interactive skills. Social networks are used by 18.6%, reflecting an interest in utilizing social platforms as educational tools; however, this usage remains low compared to videos and PowerPoint. This data indicates that while there is a trend toward using traditional interactive media, newer educational tools are still underutilized, highlighting potential areas for technological integration and teacher training.

The survey results show that media technology is mainly used by teachers for engaging students in interactive activities and presenting new content. However, its use for assessing student understanding and facilitating group collaboration is less frequent, suggesting underutilization in these areas. Key findings include: 57.2% of teachers use technology for interactive activities, 47.8% for presenting new content, 43.8% for providing additional resources, 36.2% for group work, and 16.6% for assessments. These

Table 1. Types of media technology most often used during teaching.

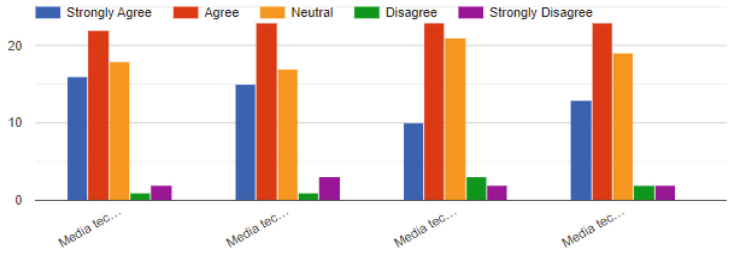


results indicate that while technology is effectively used for engagement and content delivery, there is potential to expand its role in collaboration and assessment to enhance its educational impact.

The survey indicates that teachers hold varied perceptions about the benefits of media technology in teaching. Here's a breakdown of their responses on specific aspects: Media technology enhances student engagement Table 2: 38% of teachers agree (strongly agree), suggesting that a significant portion of teachers feel media technology supports student engagement. 18% remain neutral, while only 3% disagree. Media technology improves students' understanding of complex concepts: 38% agree that technology aids in understanding complex concepts, with 17% neutral. Only a small percentage (4%) disagree, reflecting confidence in its educational value. Media technology facilitates differentiated learning: 33% of teachers agree that technology supports differentiated learning, though a higher 21% are neutral, indicating some reservations about its effectiveness here. A small 5% disagree. Media technology enhances student collaboration: 41% agree that technology promotes collaboration among students, while 19% are neutral, and only 4% disagree. This suggests that most teachers view technology as a tool to foster student interaction. Media technology makes teaching more efficient: 40% of teachers believe technology makes teaching more efficient, 15% remain neutral, and only 4% disagree. This points to a positive trend toward using technology to enhance teaching efficiency. In summary, most teachers have a favorable view of media technology in the classroom, though a considerable portion remain neutral, suggesting room for further exploration or improvement in the application of these tools.

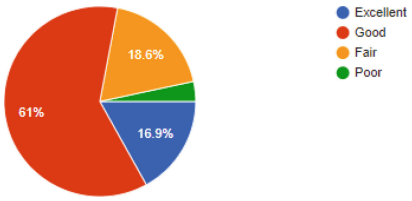
The survey shows (Table 3) that 86.3% of teachers face technical problems when using media technology in the classroom, with only a small percentage reporting no problem. This suggests that technical challenges are common and potentially disruptive to teaching. In addition, 13.7% of teacher's experience frequent technical difficulties, highlighting ongoing problems. While 64.5% of teachers have received formal training in the use of technology, 35.5% have not, suggesting that a significant proportion of teachers lack the necessary training, which could hinder the effective integration of technology in the classroom. The survey shows that the majority of teachers (77.9%) rated the training they received as 'good' or 'excellent', indicating a largely positive response to the training provided. However, a small proportion (18.6%) felt that the training was only

Table 2. Types of media technology most often used during teaching.



fair and very few felt that it was poor. This suggests that while the training programmes are largely effective, there is still room for improvement, particularly in making them more responsive to the needs of all teachers. 16.9% of teachers rate the training as 'excellent', indicating that a relatively small group had a very positive experience and found the training very useful for their needs. 61% rate the training as 'good', which is the largest percentage, suggesting that most teachers find the training effective and useful, but not outstanding. 18.6% rated the training as 'acceptable', indicating that for this group the training was useful in some respects, but had some shortcomings which limited its effectiveness. 0.5% of teachers rated the training as 'poor', indicating that only a very small percentage of teachers were dissatisfied with the training and found it ineffective.

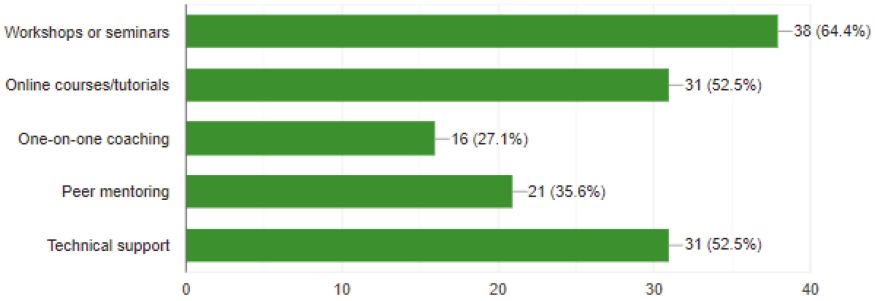
Table 3. Types of media technology most often used during teaching.



The survey data indicates (Table 4) that the majority of teachers prefer organized formats such as workshops and seminars (64.4%) and show a strong interest in online courses or tutorials (52.5%). In addition, technical support remains an important priority. Teachers are also open to mentoring colleagues, and some seek personalized support, suggesting a need for a diverse approach to meet their training needs.

In this survey, 30% of the teachers questioned belong to these non-public schools. 74% of them answered that the classrooms where they teach are equipped with smart-boards and have regular internet. They use a combination of both traditional and digital explanations. Regarding the use of media technology, they were regularly trained before the school year began. A significant 76.5% of teachers felt comfortable using media technology during lessons, while 17.3% expressed that they felt moderately comfortable, as they needed more training and occasionally required support from an expert.

Table 4. Types of media technology most often used during teaching.



6.2% of teachers stated that they did not feel comfortable using media technology in teaching.

4.2 Students, and Use of Media Technology in Teaching

20% of the students responded that teachers use smart boards equipped with internet in their classes; these students belonged to non-public schools. Meanwhile, 80% of the students said that teachers used traditional blackboards for teaching. Of the students whose teachers used smart boards, 97% said that teaching with smart boards was more effective than teaching with traditional blackboards, while only 0.3% thought that traditional teaching methods were more effective. According to the students, other teaching tools used by teachers include videos and photos, which are more commonly used in non-public schools in urban areas where classrooms are also equipped with smart boards. These tools are rarely used in public schools and almost never in public schools in rural areas. 5% of students in non-public schools said that, under the guidance of their teachers, they take quizzes at home or on weekends during study hours and participate in competitions and contests with each other using various apps and working together online.

5 Findings

There is a significant gap between private and public schools in using technology in the classroom. The survey showed that private schools have managed to integrate media technology into the learning process earlier and more systematically than public schools, which are still lagging behind. This difference clearly reflects the lack of infrastructure and support for technology in public education. Private schools, which are more focused on the quality of teaching and the use of technology to increase student engagement, have been successful in providing technological tools such as smart boards, tablets and other projects that help teachers make their teaching more effective and interactive. On the other hand, public schools, especially those in the rural areas, are still very limited in their use of technology and often lack the necessary tools to support the learning process. In urban areas, where schools are better equipped with infrastructure and technological

resources, the use of technology is more widespread and effective. 64.4% of schools in urban areas have projectors and smart boards, while this percentage is significantly lower in rural areas. This creates a significant gap in opportunities for students and teachers to fully exploit the potential of technology to improve teaching and learning. To bridge these differences and ensure that all students have equal opportunities in using technology, clear policies and investment strategies are needed from the Ministry of Education and the Albanian government. First, projects should be developed and implemented to provide public schools, especially those in rural areas, with the necessary technological infrastructure. This includes the provision of computers, tablets, high-speed internet and tools to create interactive learning environments. In addition, teachers should be provided with ongoing training to enable them to use media technology effectively in the classroom. The survey revealed that 35.5% of teachers have not received proper training in using technology, indicating that the commitment to teachers' professional development is still insufficient. In this regard, the Ministry needs to ensure support for the development of courses and seminars that will enable teachers to acquire the necessary skills to use technology as an effective teaching tool. According to another UNDP survey [35], 65% of a sample of 785 teachers from 43 schools in Albania admitted that they do not use technology, even though the curriculum allows for its integration into the learning process. This is due to the lack of computers and Internet access.

In conclusion, the gap between private and public schools in using technology needs to be closed through a strong commitment from policy makers and educational institutions. This requires sustained investment, advanced teacher training and better infrastructure in public schools, especially in rural areas, to ensure that all students have equal opportunities to use technology in education [36].

6 Discussion

Before embarking on institutional initiatives to integrate media technology into the classroom, it is essential to carefully consider both its positive and negative aspects. This critical step ensures that the technology is used effectively, maximising its benefits for teaching and learning, while minimising potential problems and negative consequences that may arise during its implementation.

A well-researched approach allows educators and policymakers to understand the true impact of media technology in the classroom. While technological tools can increase student engagement, facilitate interactive learning and provide access to a wide range of educational resources, they also pose challenges such as screen dependency, distraction, inequitable access and the need for teacher training. Through thorough research and pilot programmes, educational institutions can develop best practices that meet pedagogical goals. Investing in teacher training, ensuring equitable access to technology tools, and establishing guidelines for effective technology integration are critical steps in creating a balanced and effective digital learning environment. Ultimately, the goal should be to harness the power of media technology in a way that enhances educational quality, promotes student success and proactively addresses potential drawbacks. Only through careful planning and strategic implementation can technology truly serve as a powerful asset in modern education.

6.1 Positive Aspects

Media **technology** is merely a tool that takes the meaning through the purpose for which it is used. *"But there is a world of difference between what computers can do and what society chooses to do with them,"* said Papert & Harel [28]. The use of technology in education has many positive aspects that enhance the quality and efficiency of the teaching and learning process [33]. The use of multimedia aids in the understanding, interpretation, production and exploration of the concepts being taught. Mayer [22] noted that "*people learn more deeply when ideas are expressed in words and pictures rather than in words alone*". Technology can help to create personalised learning plans for students, taking into account their learning styles and pace. Online platforms offer opportunities for tailored teachings. When technology-based learning is combined with traditional teaching, student outcomes are significantly positive [7]. Through the Internet and other technological resources, students can access a vast amount of information and educational materials that may be more up-to-date and more diverse than traditional textbooks. Technology can facilitate collaboration between students and teachers in more creative and useful ways, such as through online communication platforms, collaborative projects and virtual discussions. The use of technology in the classroom helps students to develop their technological skills, which are essential for the job market and everyday life. Technology streamlines the assessment process in real time, allowing for faster and more accurate feedback. This helps teachers to adapt their teaching strategies to the needs of their students [13].

While the benefits of using technology are many, careful balance is needed to ensure that technology is used as a supportive tool rather than a replacement for direct teacher-student interactions. It may seem that only pencil and paper can influence children's socialisation, but the use of media technology has been shown to promote collaboration and teamwork among students (23). Online learning platforms and social media create opportunities for discussion and collaboration between students, encouraging group work and the exchange of ideas, [17]. Studies have also shown that the integration of media technology reduces students' stress and anxiety during learning [4]. Media technologies, including interactive applications and personalised learning platforms, can provide a more relaxed and manageable environment for students, reducing stress and anxiety associated with [29]; [8]. Digital storytelling as a media technology tool promotes social presence in online education, reducing student's anxiety and stress by encouraging engagement and personal connection [21]. Educational video games can also reduce stress and anxiety by providing cognitive distraction, fostering positive emotions, and improving problem-solving skills [15].

However, despite the willingness to use media technology in the classroom, several factors influence its realisation, including politics, economics, psychology, education, society and culture [34]. Studies have shown that media technology supports the development of critical thinking and problem-solving skills [19]. Tools such as simulations, educational games and discussion platforms encourage students to cultivate critical thinking and problem-solving skills through practical and collaborative activities. "*Based on this reasoning, the aim is to transform traditional classrooms into ones that are based on*

technology. This process requires teachers to be equipped with the necessary technological skills and to learn how to use new technological tools that support the teaching process and students' progress" said Hajdari, B., [16].

6.2 Negative Aspects of Technology Use

Students, parents and teachers need to be informed about the negative aspects of technology use. As students spend long hours online, there is a significant risk that this may interfere the essential time for sleep, relaxation, outdoor activities, and socialising with peers [5, 25]. Technology in education can encourage sedentary lifestyle among students, reducing opportunities for physical activity and motor skill development, which is particularly important for children in their early developmental years [6]. Teachers and schools need to teach students on critical skills [24]. These include avoiding plagiarism and being able to select relevant information from the vast and complex data available online [27].

Excessive use of technology can create dependency, making students less able to learn and think critically without technological tools [12]. Students may struggle to solve problems or learn without a device [32]. Studies have shown that the use of phones, apps and social networks for non-educational purposes reduces concentration and productivity in learning [37].

The use of different technological platforms and applications can be associated with risks related to the privacy and security of students' and teachers' data, especially if these platforms are not adequately protected. Often these issues do not appear to be managed effectively. This requires commitment from both parents and teachers. Many teachers do not use technology as a teaching system, nor do they integrate it into their [18].

7 Conclusion

The study on the use of media technology in education has highlighted some key benefits and challenges that need to be considered by educational authorities and responsible institutions. In particular, there is a noticeable gap between private and public schools in terms of technology use, as well as a significant difference between schools in urban and rural areas. This underlines the need for significant investment in educational infrastructure and teacher training to ensure that technology is used effectively and brings sustainable benefits to the teaching process.

When properly planned and implemented, the use of technology can offer significant benefits, including improved student engagement, personalised learning opportunities and the development of essential skills for the future. However, challenges such as lack of infrastructure and technological uncertainty must be taken seriously and carefully addressed to avoid negative effects such as screen dependency and loss of social interaction between students and teachers.

Policies, strategies and investments from the Ministry of Education and the Albanian government are essential to bridge these gaps and ensure the successful implementation of technology in education. These measures will ensure that every school, regardless of its location, has equal opportunities to use technology efficiently. This will contribute to

a higher quality and more accessible education for all students, preparing them for the challenges and opportunities of an increasingly digitalised world.

References

1. Albanian Ministry of Education: Strategjia Kombëtare e Arsimit 2021–2026. (National Education Strategy 2021–2026), access on web, (2020). https://arsimi.gov.al/wp-content/uploads/2021/05/Draft-Strategjia-per-Arsimin-2021-2026.pdf
2. Axhami, L., Axhami, M.: Use of Information and Communication Technologies as a Medium for Education in Albania. Int. J. Inform. Educ. Technol. 13(12) (2023)
3. Babameto, E., Pano, D.: The Future Higher Education in Albania in the Digital Era-Challenges and Opportunities. J. Electr. Syst. **20**(4s), 1809–1820 (2024)
4. Beard, C., Wilson, J.P.: Experiential Learning: A Best Practice Handbook for Educators and Trainers. Kogan Page, London (2006)
5. Biddle, S.J.H., Pearson, N., Ross, G.M., Braithwaite, R.: Tracking of sedentary behaviours of young people: A systematic revieë. Prev. Med. **51**(5), 345–351 (2010)
6. Bucksch, J., Inchley, J., Hamrik, Z., Finne, E., Tynjälä, J., Kolip, P.: Trends in ad7. lescent screen-time behaviors from 2002 to 2010 in 33 European countries. Journal of Public Health **22**(3), 179–186 (2014)
7. Christensen, C.M., Horn, M.B., Johnson, C.W.: Disrupting Class: How Disruptive Innovation Will Change the Way the World Learns. McGraw-Hill, US (2011)
8. de la Fuente, J., et al.: The interaction between self-regulation, learning approaches, resilience, and test anxiety in undergraduate students: A structural equation model. Front. Psychol. **11**, 491 (2020)
9. Dhimitri, J., Pepkolaj, L., Avdia, B.: Geography and math teachers in distance learning education amid COVID-19 pandemic in Albania. Dela **56**, 175–191 (2021)
10. Dosja, et al.: Aftësitë dixhitale, raporti i OECD: Shqipëria larg BE dhe e fundit në rajon, (Digital skills, OECD report: Albania far from the EU and last in the region) (2024).eb https://dosja.al/aktualitet/aftesite-dixhitale-raporti-i-oecd-shqiperia-larg-be-dhe-e-fundit-ne-r-i392501
11. Euronews: al, 65% e mësuesve nuk e përdorin teknologjinë në mësimdhënie, 65% of teachers do not use technology in teaching, (2024). https://euronews.al/anketa-mesuesit-nuk-perdorin-teknologjine-65-nuk-kane-akses/
12. Flanigan, A.E., Babchuk, W.A.: Social media as academic distraction: Exploring Facebook and YouTube as predictors of academic performance. J. Educ. Comput. Res. **59**(1), 152–170 (2021)
13. Garrison, D.R., Anderson, T.: E-Learning in the 21st Century: A Frameëork for Research and Practice. Routledge, UK (2003)
14. Gonzalez, C.: The Role of Blended Learning in the World of Technology. Int. J. Instr. Technol. Dist. Learn. (2004)
15. Granic, I., Lobel, A., Engels, R.C.: The benefits of playing video games. Am. Psychol. **69**(1), 66–78 (2014)
16. Hajdari, B.: Internet and technology in education, (Interneti dhe teknologjia në mësimdhënie). J. Adv. Res. Bus. Manag. Account. **5**(12), 2456–3544 (2019). https://doi.org/10.53555/nnbma.v5i12.816
17. Hamam, D., Hysaj, A.: Technological pedagogical and content knowledge (TPACK): Higher education teachers' perspectives on the use of TPACK in online academic writing classes. In: HCI International 2021-Posters: 23rd HCI International Conference, HCII 2021, Virtual Event, Proceedings, Part III 23 2021, pp. 51–58. Springer International Publishing (2021)

18. Heemskerk, I., Volman, M., Admiraal, W., ten Dam, G.: Inclusiveness of ICT in secondary education: students' appreciation of ICT tools. Int. J. Incl. Educ. **16**(2), 155–170 (2011). https://doi.org/10.1080/13603111003674560
19. Hysaj, A., Hamam, D.: Understanding the development of critical thinking through classroom debates and online discussion forums: a case of higher education in the UAE. J. Asia TEFL. **18**(1), 373–379 (2021)
20. Kraja, P.: Report on Smart Education in Albania. In: Zhuang, R., et al. Smart Education in China and Central & Eastern European Countries. Lecture Notes in Educational Technology. Springer, Singapore. (2023). https://doi.org/10.1007/978-981-19-7319-2_3
21. Lowenthal, P.R., Dunlap, J.C.: From pixel on a screen to real person in your students' lives: Establishing social presence using digital storytelling. The Internet High. Educ. **13**(1–2), 70–72 (2010)
22. Mayer, R.E.: Multimedia Learning. Cambridge University Press, Cambridge (2009). https://doi.org/10.1017/CBO9780511811678
23. McKnight, K., O'Malley, K., Ruzic, R., Horsley, M.K., Franey, J.J., Bassett, K.: Teaching in a digital age: Hoë educators use technology to improve student learning. J. Res. Technol. Educ. **48**(3), 194–211 (2016)
24. McGowan, S.: Plagiarism detection and prevention. J. High. Educ. Policy Manag. **27**(3), 319–326 (2005)
25. Ng, K.W., Augustine, L., Inchley, J.: Comparisons in Screen-Time Behaviours among Adolescents with and without Long-Term Illnesses or Disabilities: Results from 2013/14 HBSC Study. Int. J. Environ. Res. Public Health. **15**(10), 2276 (2018). https://doi.org/10.3390/ijerph15102276
26. OECD, PISA Results 2022 (Volume III) - Factsheets: Albania, (2022), web: https://www.oecd.org/en/publications/pisa-results-2022-volume-iii-factsheets_041a90f1-en/albania_88647eca-en.html
27. Park, C.: In other (people's) words: Plagiarism by university students—literature and lessons. Assess. Eval. High. Educ. **28**(5), 471–488 (2003)
28. Papert S., Harel, I.: Situating constructionism. In: Harel, I. (ed.), Constructionist learning. MIT Media Laboratory, Cambridge, MA (1990). http://www.papert.org/articles/SituatingConstructionism.html
29. Pereira, L., Rodrigues, L.: The impact of educational technology on student engagement and anxiety: A case study using mobile devices in higher education. Educ. Inf. Technol. **25**(4), 3687–3702 (2020)
30. Pepkolaj, L., Dhimitri, J.: Advancing education in albania towards effective ict integration, teacher preparation, and future-ready classrooms. Educ. J. Educ. Res. **5**(9–10), 99–109 (2023)
31. Pepkolaj, L., Duraj, S., Gerbeti, D.: The third millennium teachers train on their really needs: The Albania case. In: Conference Proceedings. The Future of Education. (2020)
32. Rideout, V.J., Foehr, U.G., Roberts, D.F.: Generation M2: Media in the Lives of 8- to 18-Year-Olds. Kaiser Family Foundation, Henry J (2010)
33. Siemens, G.: Connectivism: A learning theory for the digital age. Int. J. Instr. Technol. Distance Learn. **2**(1), 3–10 (2005)
34. Tartari, E.: Digital Technology in Education. Teknologjia Digjitale në Arsim), Publisher Luis Print, Durrës, Albania (2020)
35. UNDP Albania: Report on Training Need Assessment of Teachers of Primary and Basic Education in IT and Digital Literacy, (2024). https://www.undp.org/albania/publications/report-training-need-assessment-teachers-primary-and-basic-education-it-and-digital-literacy
36. UNESCO: Q&A: How UNESCO is driving digital learning and the transformation of education, (2024). https://www.unesco.org/en/articles/qa-how-unesco-driving-digital-learning-and-transformation-education

37. Wegmann, E., Stodt, B., Brand, M.: Addictive use of social netëorking sites and its relationship to social anxiety and loneliness. Cyberpsychol. Behav. Soc. Netw. **18**(6), 318–323 (2015)
38. Xhomara, N.: Overview of Digital Inclusion Initiatives and Perspectives in Albania. In: Tomczyk, Ł., Guillén-Gámez, F.D., Ruiz-Palmero, J., Habibi, A. (eds.) From Digital Divide to Digital Inclusion. Lecture Notes in Educational Technology. Springer, Singapore. (2023). https://doi.org/10.1007/978-981-99-7645-4_2
39. Xhumari, E., Maxhelaku, S., Como, A., nShkreli, O.: A Pupil-Centric Analysis in the Teaching-Learning Process of Information Technologies Using ICT in Albanian Schools. Br. J. Educ. **12**(4), 14–18 (2024)

Undergraduate Students' Journey with AI in the United Arab Emirates

Georgina Farouqa[1](✉), Ajrina Hysaj[2], Zaydoon Hatamleh[3], Amjad Owais[4], Othman Ibrahim[3], Laith Hiasat[5], and Sara Azeem Khan[2]

[1] Emirates Aviaition University, Dubai, UAE
georgina.farouqa@eau.ac.ae
[2] University of Wollongong in Dubai, Dubai, UAE
ajrinahysaj@uowdubai.ac.ae, sara.azeem.13@gmail.com
[3] Universiti Teknologi Malaysia, Johor Bahru, Malaysia
Zaydoon.hatamleh@gmail.com, Othmanibrahim@utm.my
[4] Al Ain University, Al Ain, UAE
amjadowais@gmail.com
[5] Rochester Institute of Technology – Dubai, Dubai, UAE
laithhiasat@gmail.com

Abstract. To maximise the benefit and minimise the risks of AI use is higher education, it is important to regulate it by institutional policies. Lack of policies and governance of AI use in higher education can hinder the use of AI tools. The research questions addressed undergraduate university students' reasons for using AI and the potential consequences that they can face because of this conduct. Qualitative research was chosen for the study, and data were collected through interviews and open-ended online questionnaire from 50 multicultural undergraduate students in the UAE, of which 10 participants were randomly selected for an in-depth interview. Resulting from thematic analysis, the findings suggested that students use AI tools for their efficiency, enhancing critical thinking, and developing writing skills. Others found AI tools costly, robotic, and unreliable. The findings also showed that students face various consequences like failing an assessment, a course, or get expelled from university. The findings potentially contribute to a better understanding of the ways that undergraduate students use AI tools in their learning process, suggesting ways that universities and academic leaders can provide resources and support for students to empower them to use AI in a creative and responsible way that can prepare students for their future roles where the use of AI can be inevitable. Universities can empower students by providing comprehensive support and resources to use AI responsibly and creatively preparing them for a future where AI will play a transformative role in various industries.

Keywords: Artificial Intelligence · Plagiarism · Efficiency · Brainstorming

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 155–171, 2025.
https://doi.org/10.1007/978-3-031-93539-8_11

1 Introduction

Gone are the days when artificial intelligence (AI) was only relevant to geeks and programmers and when students were not interested in exploring its uses [1, 32, 33, 44]. The extend by which AI use has spread throughout the world in one way or the other affects us all [2, 3, 36]. The utilization of technology by us all and especially more recently the extensive use of AI tools by students in higher education has created the necessity to explore these tools from pedagogical and holistic perspectives aiming to understand the tools and support undergraduates' students learning process through the ethical use of AI tools [4, 45, 46]. Undergraduates depend on the use of AI tools for their everyday tasks within and outside the academia; therefore, exploring the tools is not only necessary but most importantly useful and purposeful [5, 6, 30, 31]. Although the creation process of AI tools and continuous development are no doubt the expertise and contribution of coders and programmers, the utilization of GenAI tools is benefiting us all and a lot of it is even offered for free like is the case [7, 8, 18]. Undergraduate students use the AI tools to explore the topics that need to write about and often generate content which is generic and requires their understanding of the matter to become more specific, relevant and reliable [9, 13, 23, 25, 26]. Moreover, undergraduate students find academic writing complex, dry, and abstract and view AI tools as lifesavers because they provide them with the much needed "support" to produce the kind of work needed to pass subjects or even achieve a decent grade point average (GPA) [13, 14, 16].

Safeguarding academic integrity in the era of the rapid use of GenAI tools seems to be a necessity that requires the collaborations between all the stakeholders involved in the teaching and learning process such as students, educators, management of educational institutions and policy makers [15, 25, 29]. For instance, in the lenses of educators and policy makers GenAI tools, if used to deceive the whole system of higher education, are morally and academically inappropriate and deceitful [1, 19, 22]. Therefore, it is valuable that the use of GenAI tools is guided by policies put in place in governmental and institutional levels. Streamlining of the policies regarding the ethical use of GenAI tools creates a level field for all the parties involved in teaching and learning in higher education and open the path towards the ethically appropriate use of these tools in all aspects of teaching and learning such as instructional and assessment design, delivery method and in overall curriculum design [18, 32, 34, 35, 45]. These policies should take into consideration the process of developing the generically progressive set of skills that transfer from general subjects to the discipline related subjects, as well as the cumulative understanding and development of the knowledge acquisition and its application in discipline related subjects [36]. For instance, development of concepts of critical thinking, coherence, research, and inquiry are seen as crucial in all aspects of life and undergraduate students consider the empowerment through them as an open gate to better opportunities in this very globalized world, we all live in [19–21, 51]. The development of these concepts through the use of GenAI tools, serves purposefully the development of students' life skills through the exposure of students to a diverse learning environment that triggers the desire of thinking critically and analytically [22, 24, 25, 51, 57].

Utilisation of GenAI tools in higher education can be explored to increase active participation of undergraduates in the teaching and learning process, which in return can

have an impact on improving academic performance and overall university experience of undergraduate students [34]. The ethical utilization of GenAI tools can be viewed as a contributor of an added value and supporter of learning and teaching [11, 12, 23, 25, 27]. This can lead to deeper and improved understanding of the topics and improvement of the cognitive abilities related to the tasks that are utilized in the process [12, 26]. The use of GenAI in classrooms could be done on individual basis or as a group activity. In that way, students will be able to incorporate the use of GenAI to conduct research for topics of their interest and they will be subsequently familiarised with notions otherwise difficult to be explored and understood. Moreover, they will be able to approach classroom tasks with easiness and will learn and apply to persuade others and develop a sense of thought negotiation [18, 26, 27, 30]. Furthermore, the variety of topics that can be analysed through the GenAI or AI tools could be followed by research writing of individual, pairs or group reports [28, 31, 32]. Although it is a well-known fact that GenAI tools are not equally embraced and utilised by all students, yet they are powerful tools that could encourage all students to contribute equally in group work tasks without an undue external pressure imposed by the inability to complete tasks otherwise or due to the unawareness of the possible ways of resolving the tasks. Advanced writing aid tools and the tools for the effective management of the research process help students introduce novelties into their everyday study and work process and receive valuable experience in their academic achievements [9, 34, 35]. Although many studies were conducted to examine the use of AI, it was important to explore the perception of students about ways they use these tools. This study aims to explore the reasons why students in United Arab Universities use or refrain from using AI tools in their academic work as well as their awareness of potential consequences that they might face when using these tools.

2 Literature Review

2.1 Efficiency of AI Tools to Enhance Academic Work of University Students

One of the most used AI applications in education institutions is writing assistance. Analyzing writing tools, including Grammarly, Turnitin, and Hemingway Editor, enhances grammarians' checks to give students immediate feedback regarding grammar, punctuation and style. In the view of Al-Raimi et al. [4], it preserves the quality of students' written work to some extent so that students can get corrections or improvement suggestions quickly. AI tools have also changed how research is done, proving beneficial. Citations, too, can be handled automatically by reference management software, such as Zotero and EndNote. Ahmed, Siddique and Tahira [2] indicated that using such tools can significantly decrease citing time, which would otherwise be spent formatting citations and organizing the references. Also, the research process is enhanced by AI research tools such as Google Scholar and PubMed since the applications track user searches and recommend articles that may be of interest. According to Malik et al. [40], students quickly identify relevant and quality sources using artificial intelligence search engines, further enhancing their research and diversifying [37, 38].

Consequently, new generative AI tools like ChatGPT and Claude are making academic experiences among university students more remarkable by providing the best support in different areas of learning [46]. These tools are best suited to creating text

that looks like it was written by a human being and help students write better essays, reports, and research papers in terms of cohesion and lexical correctness [39, 40]. Self-designed writing tasks presented in the study have demonstrated that the expectations mentioned above are met with recent students using ChatGPT and Claude. Students' academic writing skills were better when it came to structuring the ideas and being concise [2, 41, 43]. Moreover, these tools provide an opportunity to become tutors and demonstrate how certain configuration of tools answer complicated questions or provides explanations within a specific context, which helps independent learning. Nonetheless, there is some concern that over-use of AI content generation meanwhile put the students' critical evaluation and problem-solving skills at risk [57]. Issues of ethics, like plagiarism and academic dishonesty come into the picture; therefore, one must specify use of AI solutions for learners [14, 15, 42, 44]. In general, when employing generative AI tools as a supplement to conventional methods of learning and coming up with academic results, such tools present themselves as a capacity to enhance the student's performance [29, 47].

A significant component of administrative success in university academia is the proper management of time, and AI technological assistance has greatly contributed to this aspect [4, 7]. Students can conveniently use many applications like Todoist, Trello and Notion to sort tasks, assign deadlines and manage timetables for studying [48, 49]. These tools recommend organizing the daily work depending on their deadlines and working preferences. Influenced and supported by literature, Abbas, Jam and Khan [1], found that students employing AI for time management are less likely to procrastinate and more organized. In addition, students' schedules can be scanned to provide them with lecture recommendations about feasible times to study and avoid wasting time on unimportant things and obligations. Another aspect is creating timely learning, which is also individual for every learner, thanks to AI. Interactive technologies like Coursera and Khan Academy have adopted machine learning to personalize course content based on the learner's learning abilities and rate [14, 15, 50, 53]. As Gunawardena, Bishop and Aviruppola [16] pointed out, students' learning is enhanced by adopting personalized learning since it targets areas of interest and the learning level of the students. Also, virtual tutors, now in the form of artificial intelligence, appear in education settings, particularly in universities. With the help of these virtual assistants, students can get immediate feedback, hear further explanations on some topics, or get great supplementary material to enhance their understanding of the course.

Rizvi [50] revealed that AI tutors are essential for promoting independent learning and improving students' performance, mainly in significant group learning contexts due to the lack of one-on-one teaching. Understandably, using these AI tools in academic work takes a lot of work [3, 51, 52]. Depending on the use of AI is a problem, students remain passive consumers of information made available to them. This is an apparent disregard for the crucial qualities of critical information analysis and thinking that should be fostered by students [53, 54, 57]. In their work, Ali et al. [3] stressed that it is essential that AI tools are available for all students so there are equal chances for all students to perform well in their academics. Secondly, there are ethical issues concerning data privacy, and security. Also, it is valuable to remember that many AI tools use personal data to deliver some degree of personalization, but this raises questions about whether

the data belonging to students is being adequately protected and questions regarding bias in algorithms [55, 56].

2.2 Transforming Brainstorming with AI

AI tools, such as ChatGPT, allow the possibility for the enhancement of brainstorming process. Inclusion of ChatGPT into already developed tools, such as the "kresilnik" software, allows idea generation and categorization along with the use of Open API for advanced functionalities [7, 8]. AI applied in brainstorming is a disruptive change, as it mixes traditional group ideation processes with state-of-the-art technology. While Haque et al. [20] and Li et al. [39] mention that the inclusion of AI has changed the conduct of classical brainstorming sessions, Burger et al. [7] state that AI can help generate more diverse and innovative ideas. Similarly, hybrid frameworks where human experts retain decision-making control but are supported by AI in ideation processes are promising, as underlined by Memmert and Tavanapour [42]. Furthermore, the collective intelligence, often surpassing specialised individual knowledge in areas related to radiology, has been improved through AI in a system approach; important ethical considerations, however, remain with respect to machine-generated content [16, 17, 22]. Previous studies showed what can be done in more highly complex subjects, like regional planning of development, with ideas by AI [5, 6, 38]. The emergent need for customized integration brought about developing systems that could be altered through parameters, allowing very refined outputs, such as temperature adjustment. This suggests that the temperature variation in the GPT models' parameter drives creativity, randomness, and applicability; thus, aligning with human brainstorming objectives [41].

Hence, hybrid brainstorming, with the use of multiple large language models, has emerged with promising results for enhancement in ideation and accomplishing better results through multi-system integration. Comparative analyses between the AI and human brainstorming processes yield very different entropies, latencies, and distributions. Human participants generate ideas normally distributed in length and frequency; AI operates with predefined probabilities, bound by the model's training and tokenization mechanics [47, 48]. These findings contribute to the development of information systems that can detect and distinguish between human-generated and AI-generated content, ensuring ethical and effective applications in collaborative environments [8, 14]. Brainstorming, introduced by Osborn through the book Applied Imagination [48], has gone through significant changes regarding creative problem-solving and groups generating ideas. Osborn's model was formalized to be spontaneously free but under some rules, including moving on without judgment, generating the most ideas rather than high-quality ideas in the beginning, and teamwork in developing an idea based on the work of others [48]. These provided a firm platform for brainstorming to make its application wide-ranging and varied, from even scenario planning, his work recognized some obstacles like fear of judgment and social inhibition impeding effective brainstorming.

Weiland et al. [55], noted that recently AI has emerged as the real game-changing tool to overcome these challenges. AI lets this happen by reducing the social barriers that deter participation, promotes the deeper creative explorations possible in a session, and helps maintain the anonymity of the ideas shared. AI supports individual and group ideation by providing novel starting points, synthesizing existing ideas, and catalysing

dynamic discussions. Preliminary studies have shown that AI has the potential to increase the diversity and volume of ideas in brainstorming sessions, outperforming human-only groups in certain aspects [54, 55]. These capabilities of AI also include the creation of alternative futures, the exploration of various assumptions, and refinement in archetypes of scenarios within scenario planning [52]. AI enables iteration in creating and evaluating scenarios for more depth in the exploration of uncertainties and hence enhances strategic foresight [17, 35, 52, 53]. Also, AI allows for post-session analysis by uncovering patterns and structuring outputs for actionable insights [37]. While Osborn's legacy is rooted in the creativity of the mid-20th century, his ideas fit with the developments now possible through AI. The emphasis on psychological safety and the deferral of judgment during brainstorming fits well with the ability of AI to provide a non-threatening environment for ideation [54]. Such a synthesis of Osborn's structured creativity and AI computational power provides a robust framework for modern brainstorming, especially in complex domains such as scenario planning and brainstorming.

2.3 AI Tools for Academic Writing: Potential, Challenges, and Pedagogical Integration

In the year 2017, Mizumoto et al. [43], developed Academic Word Suggestion Machine (AWSuM), which put forward frequency-based lexical bundles for particular genres. They found that in their pilot study, users valued AWSuM's suggestions, which facilitated the academic writing task and underlined the pedagogical value of this domain. Besides, AI-driven word prediction tools have shown promise to learners with disabilities. According to Evmenova et al. [10] who compared effectiveness of selected text enhancement tools in a series of studies, and evaluated tools like CoWriter and other similar tools, such as WordQ and WriteAssist the yielded increases in composition rates, spelling accuracy, and total amount of writing produced by a sub-group of students with learning incapacities. Moreover, participants enjoyed working with these tools and when contrasted with more established text enhancements, suggested significant amounts of social and educative validity [15, 49].

AI technologies have immense potential for impacting the process of writing skills improvement with different categories of learners, yet their efficiency will heavily rely on thoughtful integration into pedagogical frameworks, addressing the gaps in educators' technological knowledge, and matching tool functionalities with learners' needs. While there is research supporting the effectiveness of AI-enhanced writing tools for a short period, future research should also be directed at long-term impact on writing development, the fine-tuning of AI features to meet the demands of challenges, and a wider scope of applicability. The integration of AI into the writing curriculum should not be about technological innovation, but equally about pedagogical expertise, which will help realize the full potential of such technologies [45, 49].

3 Methodology

This qualitative research study investigates the use of Artificial Intelligence tools among undergraduate students in the United Arab Emirates and the potential consequences of their usage. The qualitative method was chosen to gain an in-depth understanding of the

students' experiences, perceptions, and challenges in using AI in their academic pursuits. Data collection was carried out in two phases: an online open-ended questionnaire and in-depth interviews. This open-ended questionnaire was designed to capture the initial insights on the motives of students in using AI tools, the perception of these tools, and the consequences they may experience. A total of 50 multicultural undergraduate students from different universities in the UAE participated in the questionnaire. The participants were selected using a purposeful sampling strategy that maximizes diversity in disciplines, cultural backgrounds, and levels of academic study. The questionnaire was sent to the participants via email and were given about 1 week to respond.

In the second phase, 10 participants were randomly selected from the questionnaire respondents for semi-structured in-depth interviews. This method enabled a deeper exploration of individual perspectives and experiences. The interviews were conducted via online platforms to facilitate the reach and flexibility of the participants who might feel constrained by the time or location. Each interview lasted approximately 30–45 min and was audio-recorded with the participants' consent. Thematic analysis was done, involving the coding of transcripts to identify key themes related to students' motivation to apply AI, experiences in using the tools, and consequences. The qualitative data were managed and analyzed with the aid of the NVivo software, making the identification of patterns and drawing of insights quite systematic and transparent. Ethical considerations are informed consent by all participants, an assurance of confidentiality, and the right to withdraw at any stage.

4 Results

The first objective of the study was to better understand the extent to which students use AI Tools in their academic work. Results are displayed in Fig. 1. The second objective of the study was to explore whether students accept or reject the use of AI tools in their academic work and reasons behind their decision. The results are displayed in Figs. 2 and 3.

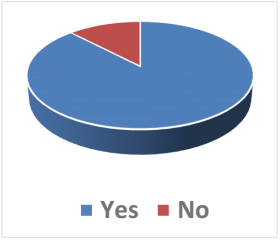


Fig. 1. Use of AI Tools

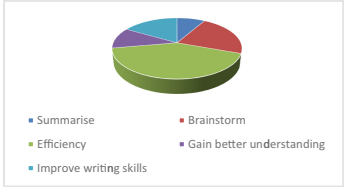


Fig. 2. Reasons for AI Use

Most of the participants used AI tools for their academic work of which 42% stated that AI tools are efficient 22% noted that AI tools help with brainstorming and organizing their thoughts. Other 17% used AI tools to improve their writing skills. Similarly, 11% of participants gained better understanding of the material using AI tools, as for improving

Fig. 3. Reasons for AI Rejection

their ability to summarize content 8% of participants mentioned that AI can help them improve this skill. The second objective of the study was to better understand the reasons why students choose to use AI tools in their academic work. The results showed that 9% of participants did not use AI tools for financial and reliability reasons. During the interviews with 10 participants, we further explored their perception of using AI tools in their academic work and reasons for use or rejection of use of AI tools. Samples of participant responses are presented in Tables 1 and 2.

The questionnaire aimed to further explore the use of AI in academic work and specifically in academic writing. The findings indicate that around half of participants use AI tools to enhance their writing skills while 33% use these tools for efficiency and 19% stated that it helps with brainstorming. As for students who refrained from using AI tools in their academic writing. More than half of students did not use AI tools because they felt they were unreliable and 42% stated that the writing outcome was robotic and lacks human touch. Results are shown in Figs. 4 and 5.

Table 1. Examples of Participant Responses

AI Use in Academic Work
Very helpful search engine, makes my work much more efficient and understanding things easier (Participant 1).
I use it to better understand concepts I am taught at university (Participant 2).
It provides me with a tool to quickly gather ideas and research (Participant 3).
Makes the work quicker (Participant 4).
I usually use it instead of Google as its more direct (Participant 5).
Sometimes, I'm on one of those days where I don't feel like using my brain, so I let AI do it for me (Participant 6).
Work smart not hard (Participant 7).

Table 2. Examples of Participant Responses

Rejection of AI Use in Academic Work
I don't believe in its reliability (Participant 8).
I don't feel like I need them. I don't think it's necessarily wrong, just not something I benefit from (Participant 9).
It is inaccurate (Participant 10).
It's stealing and even when it isn't it takes too much work for it to do what I want (Participant 8).
I wouldn't learn and I would be wasting money (Participant 8).

Fig. 4. Reasons for AI Use in Academic Writing

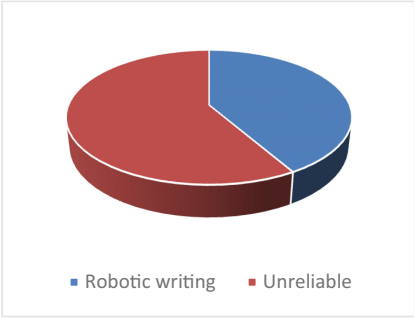


Fig. 5. Reasons for AI Rejection in Academic Writing

To gain a better understanding of the way that participants used or rejected the use of AI tools for their academic writing, we further explored this through an in-depth interview with 10 students. Samples of student responses are shown in Tables 3 and 4.

Table 3. Examples of Participant Responses

AI Use for Academic Writing
Reliable because most AI's can do their own research rather than depend on existing database (Participant 1).
You are able to gain a new academic perspective on the piece you are writing (Participant 2).
It gives phrases and words that sometimes you wouldn't have thought of, It helps me organize my thoughts to then paraphrase everything (Participant 3).
In a way it can be used to get a skeleton for a good academic piece (Participant 4).
I believe getting an outline of how the writing looks and getting some info is more efficient (Participant 5).
They save time and produce writings better than 99% of the population (Participant 6).
Well yes but I wouldn't say a very good piece. It varies (Participant 7).
It gives phrases and words that sometimes you wouldn't have thought of. It helps me organize my thoughts to then paraphrase everything (Participant 3).

In the questionnaire, 38% of the participants felt that if they use AI tools in their academic work they will be accounted for academic misconduct and practicing plagiarism. Similarly, 24% stated that they will fail the course. Some 17% were more concerned about of ethical consequences which can cause decline in creative thinking and reduction in the number of future researchers. Some 12% were confused about the policies related to the use of AI tools as they were not sure what to expect and had different feedback from instructor in this regard. A few 9% stated that if they use AI tools, they can be expelled from the university. It was important to understand how well students understand the consequences of using AI tools for their academic work. Results are shown in Fig. 6.

During interviews with participants, we were able to get a better understanding of possible consequences that they expect to face when using AI tools in their academic work. Samples of students' responses are shown in Table 5.

Table 4. Examples of Participant Responses

Rejection of AI Use in Academic Writing
If used to write a structure that will guide me, that's perfect. However, AI tools should never be used to full on write an essay (Participant 8).
Because a lot of the time they fail to use credible sources. They can make it easier by giving the writer ideas. (Participant 9).
AI on its own cannot generate a full piece of academic writing no matter the instructions given to the software. AI cannot be utilized but can be used adequately. Certain questions, public research papers, grammar and spelling, synonyms, and so on. human intervention is necessary for the most part. (Participant 10).
AI texts don't have substance and lack a human touch, it feels robotic and most of the time it feels like AI doesn't answer the question you have in mind. (Participant 8).
Because sometimes it's not as accurate (Participant 8).
While it puts ideas on the table, I never found it capable of doing the work for me (Participant 8).
It is quite simple and user friendly however, sometimes the AI system is unable to truly understand what you are asking of it which may get frustrating (Participant 8).

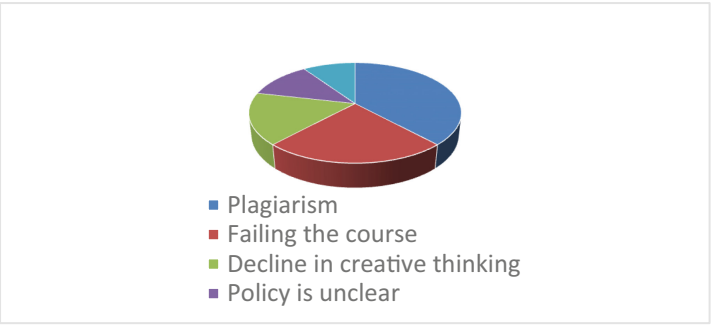


Fig. 6. Possible Consequences for AI Use

Table 5. Examples of Participant Responses

Academic Misconduct Consequences
Expulsion (Participant 1).
Plagiarism charges (Participant 1).
Charged with plagiarism consequences such as possible expulsion, removal of honours to a degree (Participant 2).
Academic misconduct, risk of disqualification (Participant 3).
AI violations typically result in dismissal from the university with an "F" grade in the course where the violation occurred and a "Withdrawal" grade in all other courses (Participant 4).
No credit for the assignment for first- or second-time offenders, failing the class or expulsion for more offenses (Participant 5).
Termination from the course if plagiarized (Participant 6).
While it puts ideas on the table, I never found it capable of doing the work for me (Participant 7).

5 Discussion

Although the use of AI by university students is common, the perception of students on reasons why they use AI and the consequences is not always explored and in most cases is under-researched, hence this study aimed at addressing this issue. Furthermore, academic institution policies about using AI are another factor that requires exploration for further clarification of policies that would benefit students' learning experience. Results showed that students use AI tools for their efficiency and ability to analyze written texts, giving immediate feedback to students regarding grammar, punctuation, and style. These results concede with studies by Al Raimi et al. [4] and Nikolic et al. [45] that viewed Al as efficient corrections, improvement suggestions, and citation tools. Similarly, Rizvi [50] revealed that AI tutors are essential for promoting independent learning and improving students' performance. Furthermore, study by Nikolic et al. [46] found that students who use AI are more efficient in time management and are less likely to delay submission of their work. Results of the current study showed that students use AI for organizing their academic writing, improving self-efficiency, and to receive timely feedback on their work, which aligns with literature by Malik et al. [40] and Ali et al. [3] which found that students identify quickly relevant and quality sources, while structuring the ideas and being concise using AI. This makes the academic experiences among university students more remarkable by providing the best support in different areas of learning [8, 9, 11, 12]. As Gunawardena, Bishop and Aviruppola [16], pointed out, students' learning is enhanced by adopting personalized learning since it targets areas of interest and the learning level of the students.

Also, the results of this study are in line with literature by Evmenova et al. [10] who compared effectiveness of selected text enhancement tools in a series of studies,

evaluating tools like CoWriter and other similar tools, such as WordQ and WriteAssist and found that students increased the understanding and application of academic writing and spelling accuracy. Furthermore, the results of this study showed that students use AI as a tool to help them brainstorm which is presented in studies by Burger et al. [7]; Haque et al. [21, 22] and Li et al. [39] that AI tools, such as ChatGPT, allow the possibility for the enhancement of brainstorming. It is important to mention that preliminary studies have also shown that AI has the potential to increase the diversity and volume of ideas in brainstorming sessions, outperforming human-only groups in certain aspects [45, 55, 56]. Some students noted that they do not use AI for their cost and lack of reliability. Students that refused to use AI in their academic writing because they felt that the written text was robotic and lacked liability. Others refused to use AI in their academic work because they felt that it would reduce productivity and the number of future researchers. These findings are in line with the study by Zhai, Wibowo and Li [57], study that showed that the overuse of AI can put students' critical evaluation and problem-solving skills at risk and crucial qualities of critical information analysis and thinking that should be fostered. The integration of AI into the writing curriculum should not be about technological innovation, but equally about pedagogical expertise, which will help realize the full potential of such technologies [49]. The study also showed that students were concerned about the consequences of using AI in their academic work; these consequences include plagiarism, failing a course or getting expelled. In addition to ethical consideration, data privacy and security as presented in the study by Ali et al. [4]. Further research is needed to explore the institutional role in setting clear policies of AI use in the academic work of university students in UAE.

6 Conclusion and Recommendations

This paper aimed at exploring the perceptions of undergraduate students of using AI tools in their academic work and possible consequences that they might face as a result of this conduct. As such, this paper aimed at contributing to better understand the ways that university students use AI tools among university without facing consequences that might result in failing courses or getting expelled because of unclear policies of AI use. The survey findings amplified the students voice about the ways they perceive the use of AI tools and their fears about the results of such use. Students should be placed at the center of the learning process, yet their views and preferences are not taken into consideration. Most students feel that the use of AI can help enhance their learning, yet they are unsure of AI use policies which leads to confusion among students. Alternatively, by using AI tools correctly, students could enhance their knowledge and become more engaged in their learning while completing their academic work. The role of educators is also crucial in achieving the best learning outcome utilizing AI as a supporting tool for the teaching and learning process. Higher education institutions need to consider the needs of students and provide them with clear policies related to AI use, aiming to prepare them for the challenges of their future professional lives, where the use of AI can be integrated as supportive and inevitable part of their daily work.

References

1. Abbas, M., Jam, F.A., Khan, T.I.: Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students. Int. J. Educ. Technol. High. Educ. **21**(1), 10 (2024). https://doi.org/10.1186/s41239-024-00444-7
2. Ahmed, M., Siddiqui, M., Usman, T.: Impact of artificial intelligence-based writing assistant on the academic writing skills of university faculty in Pakistan. Int. J. Hum. Soc. **4**(1), 539–545 (2024)
3. Ali, O., Murray, P.A., Momin, M., Dwivedi, Y.K., Malik, T.: The effects of artificial intelligence applications in educational settings: challenges and strategies. Technol. Forecast. Soc. Chang. **199**, 123076 (2023). https://doi.org/10.1016/j.techfore.2023.123076
4. Al-Raimi, M., Mudhsh, B.A., Al-Yafaei, Y., Al-Maashani, S.: Utilizing artificial intelligence tools for improving writing skills: Exploring Omani EFL learners' perspectives. Forum Linguist. Stud. **6**(2), 1177 (2024). https://doi.org/10.59400/fls.v6i2.1177
5. Aoun, J.E.: Robot-Proof: Higher Education in the Age of Artificial Intelligence. MIT Press (2017)
6. Ayoa: Ayoa Mind Mapping, Whiteboards & Tasks (2023). https://www.ayoa.com/
7. Burger, B., Kanbach, D.K., Kraus, S., Breier, M., Corvello, V.: On the use of AI-based tools like ChatGPT to support management research. Eur. J. Innov. Manage. **26**(1), 233–241 (2023). https://doi.org/10.1108/EJIM-12-2022-0665
8. Cahan, P., Treutlein, B.: A conversation with ChatGPT on the role of computational systems biology in stem cell research. Stem Cell Rep. **18**(1), 1–2 (2023). https://doi.org/10.1016/j.stemcr.2022.12.009
9. Dahri, N.A., et al.: Investigating AI-based academic support acceptance and its impact on students' performance in Malaysian and Pakistani higher education institutions. Educ. Inf. Technol. (2024). https://doi.org/10.1007/s10639-024-12599-x
10. Evmenova, A.S., Graff, H.J., Jerome, M.K., Behrmann, M.M.: Word prediction programs with phonetic spelling support: performance comparisons and impact on journal writing for students with writing difficulties. Learn. Disabil. Res. Pract. **25**(4), 170–182 (2010). https://doi.org/10.1111/j.1540-5826.2010.00315.x
11. Farouqa, G., Hysaj, A.: Exploring faculty members perception of utilizing technology to enhance student engagement in the United Arab Emirates: technology and the ICAP modes of engagement. In: International Conference on Human-Computer Interaction, pp. 67–76. Springer International Publishing, Cham (2022)
12. Farouqa, G., Hysaj, A.: Active learning in the lenses of faculty: a qualitative study in universities in the United Arab Emirates. In: International Conference on Human-Computer Interaction, pp. 77–90. Springer International Publishing, Cham (2022)
13. Foltýnek, T., et al.: Testing of support tools for plagiarism detection. Int. J. Educ. Technol. High. Educ. **17**(1), 1–31 (2020). https://link.springer.com/chapter/https://doi.org/10.1007/978-3-030-43687-2_68
14. Gao, C.A., et al.: Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. npj Digit. Med. **6**(1), 75 (2023). https://doi.org/10.1038/s41746-023-00819-6
15. Gayed, J.M., Carlon, M.K.J., Oriola, A.M., Cross, J.S.: Exploring an AI-based writing assistant's impact on English language learners. Comput. Educ. Artif. Intell. **3**, 100055 (2022). https://doi.org/10.1016/j.caeai.2022.100055
16. Gunawardena, M., Bishop, P., Aviruppola, K.: Personalized learning: the simple, the complicated, the complex and the chaotic. Teach. Teach. Educ. **139**, 104429 (2023). https://doi.org/10.1016/j.tate.2023.104429

17. Gutiérrez, L.: Artificial intelligence in language education: navigating the potential and challenges of chatbots and NLP. Res. Stud. Engl. Lang. Teach. Learn. **1**(3), 180–191 (2023). https://doi.org/10.62583/rseltl.v1i3.44
18. Hamam, D., Hysaj, A.: Technological pedagogical and content knowledge (TPACK): higher education teachers' perspectives on the use of TPACK in online academic writing classes. In: HCI International 2021-Posters: 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part III 23, pp. 51–58. Springer International Publishing (2021)
19. Hancock, J.T., Naaman, M., Levy, K.: AI-mediated communication: definition, research agenda, and ethical considerations. J. Comput. Mediat. Commun. **25**(1), 89–100 (2020)
20. Haque, A., Moon, H., Hao, H., Didari, S., Woo, J.O., Bangert, P.: Self-supervised contrastive representation learning for 3D mesh segmentation. In: Proceedings of the AAAI Conference on Artificial Intelligence (2022)
21. Haque, M.U., Dharmadasa, I., Sworna, Z.T., Rajapakse, R.N., Ahmad, H.: 'I think this is the most disruptive technology': exploring sentiments of ChatGPT early adopters using Twitter data (2022). https://arxiv.org/abs/2212.05856
22. Humble, N., Mozelius, P.: Artificial intelligence in education – a promise, a threat or a hype. In: Proceedings of the European Conference on the Impact of Artificial Intelligence and Robotics, pp. 149–156 (2019)
23. Hrytsenko, V., Tkachenko, A., Podolyan, O., Dieiev, K., Ilyn, L.: The role of artificial intelligence in personalisation of the learning process. Revista EDaPECI **24**(3), 152–165 (2024)
24. Hysaj, A., Elkhouly, A.: Why do students plagiarize? The case of multicultural students in an Australian university in the United Arab Emirates. In: ENAI Conference (2020)
25. Hysaj, A., Suleymanova, S.: Safeguarding academic integrity in crisis induced environment: a case study of Emirati engineering and IT students in a private university in the UAE. In: International Conference on Human-Computer Interaction, pp. 236–245. Springer International Publishing, Cham (2021)
26. Hysaj, A.: Group reports in the online platform: a puzzle, a ride in the park or a steep slope: a case study of multicultural undergraduate students in the United Arab Emirates. In: 2021 IEEE International Conference on Engineering, Technology & Education (TALE), pp. 745–750. IEEE (2021)
27. Hysaj, A., Hamam, D.: Understanding the development of critical thinking through classroom debates and online discussion forums: a case of higher education in the UAE. J. AsiaTEFL **18**(1), 373–379 (2021)
28. Hysaj, A., Hamam, D., Baroudi, S.: Efficacy of group work in the online platform: an exploration of multicultural undergraduates' attitudes in online academic writing classes. In: International Conference on Human-Computer Interaction, pp. 246–256. Springer International Publishing, Cham (2021)
29. Hysaj, A., Khan, Z.R.: Understanding reasons students may plagiarize in online assessments. In: European Conference on Academic Integrity and Plagiarism (2021)
30. Hysaj, A., Freeman, M., Khan, Z.R.: Theory of planned behaviour in higher education: exploring the perceptions of multicultural ESL students about cheating. In: International Conference on Human-Computer Interaction, pp. 58–71. Springer Nature Switzerland, Cham (2023)
31. Hysaj, A., Farouqa, G., Khan, S.A., Hiasat, L.: Reducing stress through formative assessments: a case of the digital platform. In: International Conference on Human-Computer Interaction, pp. 486–500. Springer Nature Switzerland, Cham (2023)
32. Hysaj, A., Farouqa, G., Khan, S.A., Hiasat, L.: A tale of academic writing using AI tools: lessons learned from multicultural undergraduate students. In: International Conference on Human-Computer Interaction, pp. 43–56. Springer Nature Switzerland, Cham (2024)

33. Hysaj, A., Freeman, M., Hamam, D.: Using AI tools to enhance academic writing and maintain academic integrity. In: International Conference on Human-Computer Interaction, pp. 57–66. Springer Nature Switzerland, Cham (2024)
34. Hysaj, A., Khan, S.A., Hamam, D., Farouqa, G.: Review of the factors that influence learners motivation in using AI-enabled gamification for collaborative learning. In: International Conference on Human-Computer Interaction, pp. 58–69. Springer, Cham (2025)
35. Jungwirth, D., Haluza, D.: AI-based scenario generation for future planning: an exploratory study using GPT-3. J. Curr. Trends Comput. Sci. Res. **2**(2) (2023). https://doi.org/10.33140/JCTCSR.02.02.02
36. Khan, Z., Sivasubramaniam, S., Anand, P., Hysaj, A.: The role e-tools play in supporting teaching and assessments with integrity during the COVID-19 pandemic. In: European Conference on Academic Integrity and Plagiarism 2021: Book of Abstracts, pp. 52–54. European Network for Academic Integrity (ENAI) (2021)
37. Kozachek, D.: Investigating the perception of the future in GPT-3, -3.5 and GPT-4. In: ACM C&C 2023: Proceedings of the 15th Conference on Creativity and Cognition, pp. 282–287 (2023). https://doi.org/10.1145/3591196.3596827
38. Lavrič, F., Škraba, A.: Brainstorming will never be the same again—a human group supported by artificial intelligence. Mach. Learn. Knowl. Extract. **5**(4), 1282–1301 (2023). https://doi.org/10.3390/make5040065
39. Li, N., Yan, Y., Yang, Y., Gu, A.: Artificial intelligence capability and organisational creativity: the role of knowledge sharing and organisational cohesion. Front. Psychol. **13**, 845277 (2022). https://doi.org/10.3389/fpsyg.2022.845277
40. Malik, A.R., et al.: Exploring artificial intelligence in academic essay: higher education student's perspective. Int. J. Educ. Res. Open **5**, 100296 (2023). https://doi.org/10.1016/j.ijedro.2023.100296
41. Marion, S.: How to Use OpenAI Model Temperature? (2023). https://gptforwork.com/guides/openai-gpt3-temperature
42. Memmert, L., Tavanapour, N.: Towards human-AI-collaboration in brainstorming: empirical insights into the perception of working with a generative AI. In: ECIS 2023 Proceedings (2023). https://aisel.aisnet.org/ecis2023_rp
43. Mizumoto, A., Hamatani, S., Imao, Y.: Applying the bundle-move connection approach to the development of an online writing support tool for research articles. Lang. Learn. (2017) https://doi.org/10.1111/lang.12250
44. Naimi, A., Yaqot, E., Aissa, H.: ChatGPT as an essential tool: student perceptions in academic writing at Chlef University. J. Stud. Lang. Cult. Soc. **7**(2), 98–114 (2024)
45. Nikolic, S., et al.: ChatGPT versus engineering education assessment: a multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity. Eur. J. Eng. Educ. **48**(4), 559–614 (2023)
46. Nikolic, S., et al.: ChatGPT, Copilot, Gemini, SciSpace and Wolfram versus higher education assessments: an updated multi-institutional study of the academic integrity impacts of Generative Artificial Intelligence (GenAI) on assessment, teaching and learning in engineering. Australas. J. Eng. Educ. **29**(2), 126–153 (2024)
47. OpenAI: Tokenizer—OpenAI Platform (2023). https://platform.openai.com/tokenizer
48. Osborn, A.F.: Applied Imagination: Principles and Procedures of Creative Problem-Solving (3rd rev. ed.). Charles Scribner's Sons (1963)
49. Powers, D.E., Burstein, J.C., Chodorow, M., Fowles, M.E., Kukich, K.: Comparing the validity of automated and human essay scoring. ETS Res. Rep. Ser. **2**, 1–23 (2000). https://doi.org/10.1002/j.2333-8504.2000.tb01833.x
50. Rizvi, M.: Investigating AI-powered tutoring systems that adapt to individual student needs, providing personalized guidance and assessments. Eurasia Proc. Educ. Soc. Sci. **31**, 67–73 (2023). https://doi.org/10.55549/epess.1381518

51. Sivasubramaniam, S., Khan, Z., Anand, P., Hysaj, A.: 'e'-thinking teaching and assessment to uphold academic integrity: lessons learned from emergency distance learning. Int. J. Educ. Integrity 17(1), 17 (2021)
52. Sohrabi, S., Riabov, A., Katz, M., Udrea, O.: An AI planning solution to scenario generation for enterprise risk management. Proc. AAAI Conf. Artif. Intell. **32**(1), 11304 (2018). https://doi.org/10.1609/aaai.v32i1.11304
53. Spaniol, M.J., Rowland, N.J.: AI-assisted scenario generation for strategic planning. Futures Foresight Sci. **5**(2), Article e148 (2023). https://doi.org/10.1002/ffo2.148
54. Van der Heijden, K.: Scenarios: The Art of Strategic Conversation. John Wiley & Sons (1996)
55. Weiland, S., Jung, D., Haluza, D.: Brainstorming with chatbots: an experimental study on human-machine collaboration. J. Creat. Res. **12**(3), 1–10 (2022). https://doi.org/10.3389/frai.2022.880673
56. Yu-Han, C., Chun-Ching, C.: Investigating the impact of generative artificial intelligence on brainstorming: a preliminary study. In: 2023 International Conference on Consumer Electronics-Taiwan (ICCE-Taiwan), pp. 193–194. IEEE (2023). https://doi.org/10.1109/ICCE-Taiwan58799.2023.10226617
57. Zhai, C., Wibowo, S., Li, L.D.: The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. Smart Learn. Environ. **11**(1), 28 (2024). https://doi.org/10.1186/s40561-024-00316-7

Perspectives of Faculty on the Easiness and Usefulness of AI Tutoring Systems in Higher Education

Doaa Hamam[1(✉)], Yasser Ouda[2], Ajrina Hysaj[3], Mahinour Ezzat[1], and Sara Azeem Khan[3]

[1] Higher Colleges of Technology, Dubai, UAE
dhamam@hct.ac.ae
[2] NHS, Northumbria, UK
[3] University of Wollongong in Dubai, Dubai, UAE

Abstract. No one can deny that AI is becoming an integral part of many aspects of the educational field. One major use of AI is the AI-enabled tutoring systems. These play an important role in students' progress and improvement levels. However, being a new concept, it was necessary to explore them further to know more about their perceived usefulness and ease of use. The focus of this study was on faculty as one of the most important stakeholders. Aiming to create a robust conceptual framework, this study used the technology acceptance model (TAM) combined with elements of the constructivism learning theory and socio-technical systems theory (STS). The study explored the faculty's perceptions about the usefulness and ease of using AI tutoring systems. The study employed a survey comprised of 15 Likert scale questions where the answers varied from strongly agree to strongly disagree, addressing the psychological, pedagogical, and organisational aspects of the use of AI tutoring systems from the perspective of Higher Education teachers in the United Arab Emirates. For that purpose, 50 teachers participated in the anonymous survey, and despite the challenges pointed out by some, most participants found the use of AI tutoring systems very useful. In addition, the study found that these tools are relatively easy to use by students whose first language is not English, with minimum faculty guidance and facilitation, while at the same time ensuring fairness and equity. This study aimed to open the path to further exploration of the use of AI tutoring systems by other key stakeholders, such as students and policymakers.

Keywords: AI Tutoring Systems · TAM Model · Socio-Technical System · Constructivism

1 Introduction

AI Tutoring Systems (AITS) are considered a relatively new concept in higher education [24]. The popularity of AITS has seen a constant rise in higher education mainly due to the possibility of improving students' active engagement, enhancing classroom experience

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 172–179, 2025.
https://doi.org/10.1007/978-3-031-93539-8_12

and subsequently improving learning outcomes [1]. Like any other form of technology, the implementation of AITS requires appropriate understanding that derives from a thorough exploration of faculty perceptions about the usefulness and ease of using these powerful tools to facilitate the process of teaching and learning in higher education. One of the aspects that need exploration and understanding is the utilisation of the capacities of these tools to create an inclusive classroom environment that encourages student's involvement in collaborative learning [24]. One of the features valued as necessary when considering the development of students' new knowledge is the ability these tools offer to create new knowledge constructed on existing knowledge. AITS tools allow the learners to experiment with their learning by brainstorming ideas or giving prompts to generate concepts to form an understanding and to inform their learning [25].

In general, AITS are platforms that allow the continuous development of students' skills while creating an understanding of the topic and the variety of ways it can be explored and analysed. While faculty are eager to use AITS to support the teaching and learning process, it is important that these tools are explored to understand their characteristics, especially in relation to their usefulness and ease of use [26]. The consideration of AITS usefulness is related to instructional, curriculum and assessment design [34], aiming to motivate learners to actively participate in the process of teaching and learning [27]. This kind of knowledge acquisition allows the utilisation of existing information while purposefully and actively constructing new knowledge [17, 34]. While many fields have found the integration of AI helpful and easy to use, the implementation of AI tutoring systems in higher education requires not only the exploration of the benefits but also the challenges connected with their implementation [8].

The current study explores several aspects of these systems, focusing on their perceived usefulness and ease of use. The literature review of the current study will look into three aspects: the integration of these systems and their usefulness and ease of use from the higher education (HE) teachers' perspective, and the focus will also be on the benefits and the challenges encountered when using the AI-enabled tutoring systems. Faculty perceptions of AITS's ease of use and usefulness in higher education teaching and learning are directly related to the encouragement of utilising these tools to increase teaching and learning efficacy and facilitate students' life-long learning [11, 12]. Their active engagement highly influences students' experience in higher education in the learning process, which translates to better academic performance [17, 19]. According to the concept of constructivism, active engagement is crucial for building on students' existing knowledge and constructing new knowledge [6]. While the constructivism theory is an important paradigm for teaching and learning when considering the process of knowledge acquisition, in the context of AITS, it becomes valuable to connect it with the theories of the Technology Acceptance Model (TAM) by Davis [9] and the socio-technical system theory by Trist [32] to form the theoretical framework of the study. To do the research on this topic, the researchers wrote the following research question: "What are the perspectives of faculty members regarding the perceived ease of use and usefulness of AITS in improving the teaching and learning processes in higher education?

2 Literature Review

According to Zhai et al. [34], using AI in education transforms teaching and learning practices. Integrating and using AI in the education field has many benefits, especially in the context of higher education [22]. AI Tutoring Systems (AITS) are defined as apps or platforms that utilise AI to help and support students in their studies [33]. Currently, a plethora of educational tools and platforms can guide and aid students in progressing in their studies [31]. If we look at AI-enabled tutoring systems, we will find that they play a crucial role in students' comprehension and advancement because of their outstanding features, like instant feedback [7] and customised learning. They also provide real-time analysis of the student's progress and a prediction for their expected levels and grades [10]. These benefits are aligned with the socio-technical system theory [31].

2.1 Faculty Perceptions of AI-Enabled Tutoring Systems

It is essential to know about higher education teachers' perspectives to ensure the success of AI-enabled systems and their use in higher education. For this purpose, we used the Technology Acceptance Model (TAM) [15]. Higher education teachers can integrate AITS into their practice to help support their teaching and learning goals. However, HE teachers usually build their expectations based on their previous interactions with technology and the training they received [21]. In a study by Al-Mughairi [4], teachers expressed that they are satisfied with the benefits offered by AI platforms such as ChatGPT and mentioned that using this platform saves time and effort. However, they were also worried about potential security threats and data privacy. When it comes to the implementation part, Kaplan-Rakowski, et al. [19] mentioned that students were actively engaged in the interaction with AI.

In the UAE, the students' first language is Arabic, and English is considered their second language [18]. Therefore, there are a lot of benefits of using AITS in improving the students' second language levels, not only in colleges and universities but also in the workplace. According to Almelhes [3]: "organizations could utilize AI technology to develop their workers' knowledge; individual learners could use AI technology to facilitate their studies anywhere and anytime; and traditional learning institutions could incorporate AI-powered language learning methods to diversify learners' opportunities. There are many benefits to employing AI in language learning, particularly in second-language learning (p. 1259)."

2.2 AITS: Challenges and Benefits

Although HE teachers express many benefits for AITS, some challenges do exist. Technical issues, ethical considerations, and the potential of relying solely on AI without enough human interaction, like group work [16], were other challenges mentioned by faculty in the existing literature [5, 14]. Another issue is the lack of governance and policies for the new emerging use of AI in education, although institutions are working fast to provide these. Some other challenges can be the threat to the security and privacy of users and the threat to sustainability, according to Okulich-Kazarin et al. [28]. On the other hand, the benefits of AITS are too many. Kim [21] mentioned that AI educational

tools help elevate science and technology knowledge; however, professional development plays an important role in preparing teachers for this new era. In general, the use of AI is generally accepted and perceived as applicable in the HE education field.

3 Methods

This study used the Technology Acceptance Model (TAM), constructivism learning theory, and socio-technical systems theory (STS) to examine faculty perceptions of AI-enabled tutoring systems in higher education. The research was done within the context of the UAE, focusing on the psychological, pedagogical, and organisational aspects of AI tutoring systems. A quantitative method was employed, using a survey of 15 Likert scale questions, with the options strongly agree, agree, disagree, strongly disagree. The survey questions were designed to explore the faculty's perceptions of the usefulness and ease of use of AI tutoring systems, emphasising their role in facilitating learning and ensuring fair practices for students whose first language is Arabic. The survey was given to a purposeful sample of 50 higher education teachers from several universities in the UAE. The participants were told that their participation would be voluntary, that their personal data would not be revealed to any third party, and that their identities would remain anonymous. The participants were also told that they could withdraw from the study at any time without the need for any justification. The collected data was meant to explore the teachers' perspectives by addressing different aspects of the AI tutoring system implementation, including teachers' roles in advising students and the perceived benefits for Arabic speakers who use this system in their second language, English. There was also a section about the organisations' methods of implementation and use. The focus of data analysis was to identify the teachers' perceptions and, at the same time, identify challenges and benefits that are relevant to the AI tutoring systems.

4 Results

The survey responses showed a positive perception of the usefulness of AI tutoring systems. 85% of the participants either Strongly Agreed or Agreed that AI tutoring systems contributed to better learning outcomes. Also, 83% of respondents thought that the AI tutoring systems make it easier for students to understand complex concepts in a more efficient way compared to the traditional methods. The next item was about the impact on the overall academic environment, 87% of participants expressed their agreement that AI tutoring systems could improve students' performance, with only 6% expressing disagreement. In addition, 85% of respondents stated that AI tutoring systems direct teachers to use their teaching strategies more effectively. The next topic was the ease of use of AI tutoring systems, which was also agreed upon by the participants, 82% of respondents mentioned that AI tutoring systems are easy to use, even if they do not have enough knowledge of the technical issues, while 85% indicated that knowing how to use the systems took some time and effort. 83% of the respondents expressed that they received enough training and support to use the AI tutoring systems. Moreover, 85% of respondents expressed that the design of AI tutoring systems gave easy access to students who come from different backgrounds. Furthermore, the AI tutoring systems

were praised because they helped the teachers in facilitating learning. Most respondents (90%) mentioned that AI tutoring systems gave more opportunities for personalised learning as they addressed the students' specific needs. Another significant issue that was discussed in this study was the issue of the given feedback. 85% of respondents described the given feedback as immediate and useful. Also, 83% of participants thought that AI tutoring systems improve critical thinking and problem-solving skills, while 85% expressed the effectiveness of the systems in encouraging and improving engagement in the given learning activities. Finally, the respondents' views about the alignment of AI tutoring systems were generally positive, with 80% of participants mentioning that their institutions had the required infrastructure to use the AI tutoring systems effectively. 85% of participants agreed that AI tutoring systems work well with higher education practices regarding teaching and learning.

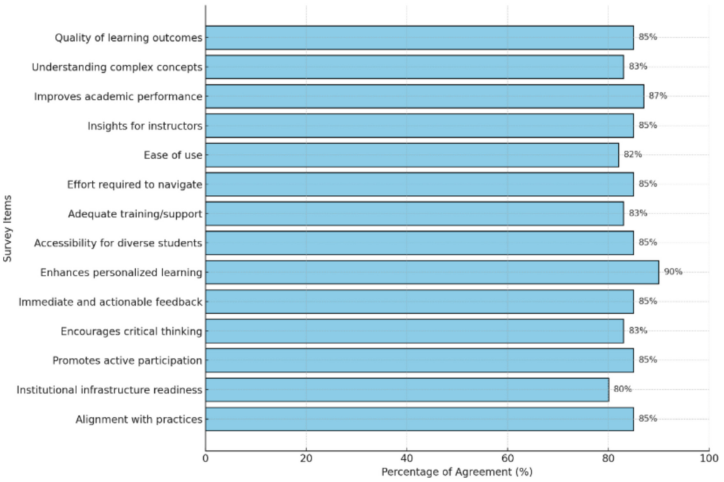


Fig. 1. HE Teachers' Perspectives on the Use of AITS.

5 Discussion

This study found that the utilization of AITS in supporting students' learning in higher education is innovative yet challenging. This finding agrees with those of Farouqa and Hysaj [11] and Paul [29], who highlighted the complexity of using these tools in teaching and learning in higher education. Furthermore, this study found that adapting AITS agrees with the TAM model, improves student learning outcomes, provides personalised support, and enhances critical thinking [13]. According to this study, the faculty members agreed that using AI tutoring systems allows students to take charge of their learning, which can have a positive effect on their learning outcomes, create a more inclusive learning environment through the provision of personalised support, and, most importantly, can potentially enhance their critical thinking skills. This finding supports the findings of Granić [13] and Lai et al. [23], who emphasised the faculty perceived

the use of AITS to positively affect students' personal and academic growth. Moreover, this research highlighted the importance of purposeful and effective use of AITS, which concurs with the findings of Peng [30].

Studies by Peng et al. [30] and Paul [29] highlighted the importance of quality assurance when implementing AITS in higher education. Similarly, this study emphasised that faculty perceptions of AITS integration seek alignment of the purposefulness of this integration in the context of higher education while highlighting the need to focus on the components of quality assurance. The integration of AITS in higher education seeks to encourage faculty acknowledgement of the benefits deriving from the use of these tools while exploring the challenges and drawbacks of their implementation [2]. This study found that the most valuable aspect of AITS integration in teaching and learning in higher education is the personalization of learning, which creates the path for knowledge acquisition, creation and recreation. This agrees with the need to encourage the use of AITS while emphasising the need for constructive teaching and learning in higher education [29, 30].

6 Conclusion and Recommendations

The results of this study indicate that faculty in higher education across the United Arab Emirates have primarily a positive perspective on AITS, highlighting both their perceived effectiveness and relative ease of use. In alignment with the Technology Acceptance Model (TAM), HE teachers agreed that AI tutoring systems can improve student learning outcomes, provide personalised support, and enhance critical thinking. Most respondents emphasised how these systems simplify complex concepts, improve academic performance, and offer personalised feedback, all considered essential elements to enhance student engagement and success. Also, HE teachers confirmed that the AI tutoring systems are easy for users to use, with the majority of respondents confirming that AITS requires little technical skills. Faculty mentioned that AITS have the potential to support students whose second language is English, and it leads to a better learning experience. Future research should focus on the perspectives of various stakeholders, such as students, policymakers, and HE leaders. This focus can lead to more understanding of the entire system in which AITS work, offering more understanding of their pedagogical effect. Although many HE teachers agreed with the AITS's ease of use, they also expressed that continuous professional development and training on such systems are essential for enhancing the potential of these systems. So, PD programmes, training sessions, and efficient technical support should enable HE teachers to improve their pedagogical strategies and utilise AI for an informed teaching practice. On the other hand, the study highlights how AI tutoring systems can benefit all types of students, particularly those whose English is their second language. Therefore, the use of AITS should help promote inclusive learning opportunities for all learners. It is also essential to have an ongoing evaluation of the integration of AITS in the teaching and learning processes, and it is also essential to get regular feedback while involving students, teachers, and educational administrators when using the AITS. In addition, continuous monitoring of learning outcomes and satisfaction levels will help decision-makers address any challenges of AITS and improve the implementation strategies. In conclusion, AITS has the potential to provide personalised, efficient, and inclusive learning

environments, although some challenges have been observed. Finally, while the overall perception among faculty is highly positive, professional development and training remain essential for the overall success of the process.

References

1. Alam, A.: Harnessing the power of AI to create intelligent tutoring systems for enhanced classroom experience and improved learning outcomes. In: Intelligent Communication Technologies and Virtual Mobile Networks, pp. 571–591. Springer Nature Singapore, Singapore (2023)
2. Al-Hattami, H.M.: Understanding perceptions of academics toward technology acceptance in accounting education. Heliyon **9**(1) (2023)
3. Almelhes, S.A.: A review of artificial intelligence adoption in second-language learning. Theory Pract. Lang. Stud. **13**(5), 1259–1269 (2023)
4. Al-Mughairi, H., Bhaskar, P.: Exploring the factors affecting the adoption of AI techniques in higher education: insights from teachers' perspectives on ChatGPT. J. Res. Innov. Teach. Learn. (2024)
5. Amah, O., Okesipe, V.O.: AI in education: ethical challenges and opportunities. In: Integrating Artificial Intelligence in Education: Enhancing Teaching Practices for Future Learning, pp. 1–26 (2025)
6. Bada, S.O., Olusegun, S.: Constructivism learning theory: a paradigm for teaching and learning. J. Res. Method Educ. **5**(6), 66–70 (2015)
7. Baroudi, S., Aoun, S., Hamam, D.: Using feedforward to improve pre-service teachers' academic writing and critical thinking skills. Front. Educ. **8**, 1126594–1126599 (2023)
8. Chuang, S.: The applications of constructivist learning theory and social learning theory on adult continuous development. Perform. Improv. **60**(3), 6–14 (2021)
9. Davis, F.D.: Technology acceptance model TAM. In: Al-Suqri, M.N., Al-Aufi, A.S. (eds.) Information Seeking Behavior and Technology Adoption, pp. 219–227 (1989)
10. Ellikkal, A., Rajamohan, S.: AI-enabled personalized learning: empowering management students for improving engagement and academic performance. Vilakshan-XIMB J. Manage. (2024)
11. Farouqa, G., Hysaj, A.: Exploring faculty members' perception of utilizing technology to enhance student engagement in the United Arab Emirates: technology and the ICAP modes of engagement. In: International Conference on Human-Computer Interaction, pp. 67–76. Springer International Publishing, Cham (2022)
12. Farouqa, G., Hysaj, A.: Active learning in the lenses of faculty: a qualitative study in universities in the United Arab Emirates. In: International Conference on Human-Computer Interaction, pp. 77–90. Springer International Publishing, Cham (2022)
13. Granić, A.: Technology acceptance and adoption in education. In: Handbook of Open, Distance and Digital Education, pp. 183–197. Springer Nature Singapore, Singapore (2023)
14. Hazaimeh, M., Al-Ansi, A.M.: Model of AI acceptance in higher education: arguing teaching staff and students' perspectives. Int. J. Inf. Learn. Technol. **41**(4), 371–393 (2024)
15. Hysaj, A., Hamam, D., Baroudi, S.: Efficacy of group work in the online platform: an exploration of multicultural undergraduates' attitudes in online academic writing classes. In: International Conference on Human-Computer Interaction, pp. 246–256. Springer International Publishing, Cham (2021)
16. Hysaj, A.: COVID-19 pandemic and online teaching from the lenses of K-12 STEM teachers in Albania. In: 2021 IEEE International Conference on Engineering, Technology & Education (TALE), pp. 1–7. IEEE (2021)

17. Hysaj, A., Hamam, D.: Dimensions of formative feedback during the COVID-19 pandemic: evaluating the perceptions of undergraduates in multicultural EAP classrooms. In: International Conference on Human-Computer Interaction, pp. 103–114. Springer International Publishing, Cham (2022)
18. Hysaj, A., Farouqa, G., Khan, S.A., Hiasat, L.: Reducing stress through formative assessments: a case of the digital platform. In: International Conference on Human-Computer Interaction, pp. 486–500. Springer Nature Switzerland, Cham (2023)
19. Kaplan-Rakowski, R., Grotewold, K., Hartwick, P., Papin, K.: Generative AI and teachers' perspectives on its implementation in education. J. Interact. Learn. Res. **34**(2), 313–338 (2023)
20. Kim, J.: Leading teachers' perspective on teacher-AI collaboration in education. Educ. Inf. Technol. **29**(7), 8693–8724 (2024)
21. Kuleto, V., et al.: Exploring opportunities and challenges of artificial intelligence and machine learning in higher education institutions. Sustainability **13**(18), 10424 (2021)
22. Lai, C.Y., Cheung, K.Y., Chan, C.S.: Exploring the role of intrinsic motivation in ChatGPT adoption to support active learning: an extension of the technology acceptance model. Comput. Educ. Artif. Intell. **5**, 100178 (2023)
23. Lin, C.C., Huang, A.Y., Lu, O.H.: Artificial intelligence in intelligent tutoring systems toward sustainable education: a systematic review. Smart Learn. Environ. **10**(1), 41 (2023)
24. Maity, S., Deroy, A.: Generative AI and its impact on personalized intelligent tutoring systems (2024). arXiv:2410.10650
25. Marouf, A., Al-Dahdooh, R., Ghali, M.J.A., Mahdi, A.O., Abunasser, B.S., Abu-Naser, S.S.: Enhancing education with artificial intelligence: the role of intelligent tutoring systems (2024)
26. Narayan, R., Rodriguez, C., Araujo, J., Shaqlaih, A., Moss, G.: Constructivism—Constructivist Learning Theory (2013)
27. Okulich-Kazarin, V., Artyukhov, A., Skowron, Ł, Artyukhova, N., Wołowiec, T.: Will AI become a threat to higher education sustainability? A study of students' views. Sustainability **16**(11), 4596 (2024)
28. Paul, R.: The politics of regulating AI technologies: towards AI competition states. In: Handbook on Public Policy and Artificial Intelligence, pp. 261–279. Edward Elgar Publishing (2024)
29. Peng, M.Y.P., Xu, Y., Xu, C.: Enhancing students' English language learning via M-learning: integrating technology acceptance model and SOR model. Heliyon **9**(2), e13302 (2023)
30. Ritterbusch, G.D., Gronau, N.: Training in industry 4.0 with AI tutoring systems. Industry 4.0 Sci. **40**(5), 50–57 (2024)
31. Shahzad, U., Dong, Z.: Prerequisites of Human-AI Collaboration: Knowledge Diffusion and Management Innovation (2024)
32. Trist, E.L.: The Evolution of Socio-Technical Systems, vol. 2. Ontario Quality of Working Life Centre, Toronto (1981)
33. Wang, S., Yu, H., Hu, X., Li, J.: Participant or spectator? Comprehending the willingness of faculty to use intelligent tutoring systems in the artificial intelligence era. Br. J. Educ. Technol. **51**(5), 1657–1673 (2020)
34. Zhai, X., et al.: A review of Artificial Intelligence (AI) in education from 2010 to 2020. Complexity **1**, 8812542 (2021)

Exploring the Use of Paraphrasing Tools in Academic Writing and Its Potential Relation with Instances of Plagiarism

Ajrina Hysaj[1(✉)], Sara Azeem Khan[2], and Georgina Farouqa[3]

[1] UOWD College, University of Wollongong in Dubai, Dubai, UAE
ajrinahysaj@uowdubai.ac.ae
[2] University of Wollongong in Dubai, Dubai, UAE
[3] Emirates Aviation University, Dubai, UAE
georgina.farouqa@eau.ac.ae

Abstract. As multicultural students use paraphrasing or translation tools very often in academic writing, it is valuable to explore their utilization for the benefit of teaching and learning in our ever-changing world. This study aimed to contribute to the ongoing research on measuring multicultural students' attitudes towards plagiarism and academic writing using paraphrasing or translation tools. This study contributes to the ongoing research on measuring students' attitudes towards academic writing using technology and plagiarism. For this study, a survey was conducted with questions adapted from Razi (2015), comprising 10 Likert-scale items designed to understand how multicultural students perceived their existing needs related to academic writing, paraphrasing tools and plagiarism. The number of respondents was 206 and they at the time of the study were enrolled in two international universities in 1) Turkiye and 2) United Arab Emirates. Data was collected over a period of three months through Qualtrics and the respondents included over 60 nationalities, majority spoke English as a second language. The results illustrate students' understanding of their needs related to academic writing skills, the use of paraphrasing and translation tools and what they consider as acceptable in terms of academic misconduct. This study found that paraphrasing and translation tools are technological advances that are used by students to complete academic writing tasks, hence, it is important that educators are well aware of the benefits and the drawbacks that these tools represent for the benefit of students and educators alike. The chapter ends with recommendations for further research.

Keywords: Paraphrasing Tools · Academic Writing Skills · Plagiarism · Multiculturalism · Critical Thinking · Academic Reading

1 Introduction

Students use paraphrasing or translation tools to complete academic writing tasks, hence, these tools need to be explored and utilized, in the benefit of teaching and learning in our ever-changing world and not against it [7]. A number of studies have attempted to

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 180–193, 2025.
https://doi.org/10.1007/978-3-031-93539-8_13

scientifically explore and understand the impact of paraphrasing tools in students' ability to master academic writing skills in light of integrity [39, 42, 43]. The sudden outbreak of COVID19 confronted the concept of face-to-face classes and increased the amount and the kind of challenges faced by federal and international universities worldwide [12, 19, 26]. Furthermore, utilization of these evolving tools in higher education (HE) and in academic writing classes is now applicable in most universities worldwide [13, 22]. As these paraphrasing and translation tools are commonly used by students worldwide, over the past decades researchers and educators have been engaged in a continuous discussion about the ethical implications connected with the use of these tools. Given the severity and increased occurrence of cheating and e-cheating, some studies have suggested a variety of strategies that some academic institutions apply to address academic dishonesty [26]. Moreover, students' inability to write appropriately essays, reports or thesis has recently gained an emphasized momentum due to the widespread of distance or blended learning [16, 20].

Academic writing is a very important and functional type of writing which is used throughout university years [19, 20]. Due to the outspread of its use and the complexity that it presents, academic writing has long drawn the attention of researchers and educators worldwide [22]. Academic writing could present itself in a form of a basic academic style, general academic style, discipline specific academic style and process-solution academic style [41]. The concepts of academic writing skills can be categorized as citation, idea presentation, academic writing and mechanics [23, 24, 41]. As Bertram Gallant [2], rightly pointed out, while considering the degree and purpose of utilization of academic writing, it is somewhat utopic to imagine that its construction is merely achievable through the introduction of its components without focusing on the development and the application of them. Academic writing integrates the individual voice of the writer with those of other researchers' voices, with the ultimate determination of achieving a smooth flowing discourse of author' and external' voices [21]. As students find academic writing complex and somewhat irrelevant to their discipline related assessment tasks, they tend to depend on paraphrasing and translation tools like Grammarly or Google translate to produce academic writing [31, 34].

According to Bloom's taxonomy [3], the process of shaping our thinking behaviour is separated in two main categories; a) lower and b) higher-order thinking skills. These categories are conceptualized in a hierarchical style and the purposes they are utilized for differ. A similar model of classification was developed by Maki [34] who classified the four levels of learning as the ones involved in remembering, processing, creating, and evaluating. Furthermore, Maki mentioned that the levels of evaluating and creating are an indispensable part in the development and application of critical and analytical thinking in academic writing classes. Critical and analytical thinking skills has been researched extensively for the last three decades with authors like McKinley [35] and McPeck [36] analyzing the uses of critical and analytical thinking in social or academic settings. Processes of development and application of academic writing skills necessitate the existence of critical and analytical thinking during academic reading and academic writing processes [15, 19, 26].

As students are partners in teaching and learning, there is a growing need to adopt a holistic approach to address the issue of the use of technology and its correlation with

development of academic writing and implication on instances of academic dishonesty among students in universities [1, 42]. It is essential that educators explore ways of helping students master their academic writing skills without compromising on quality, integrity or institutional culture. The empirical data on cheating has recently become even more alarming, with studies by Khalmetski and Sliwka [31] and Fendler, Yates and Godbey [10] and Rogerson [46] that highlight the need of understanding the effects of paraphrasing or translation tools usage in HE especially in production of written work. This study aimed at exploring factors and ways related to the use of paraphrasing and translations tools by students when completing tasks such as academic writing. It is important to highlight that majority of students who took part in this study had previously used these specific tools so no prior induction to the tools was required. The main questions that the authors of this study aimed to explore and analyse were as follows:

a) Do students use paraphrasing and translation tools when completing academic writing tasks?
b) Do students breach academic integrity when using paraphrasing and translation tools?

2 Literature Review

2.1 Delineating Academic Writing Through Development of Critical Thinking and Formative Feedback

Academic writing is commonly considered by students as a very complex, tedious, unnecessary and time-consuming process [14, 29, 34, 46]. In 2019, Strobl et al. [47] stated that academic writing is somehow mysterious and its practices and their application are poorly understood by teachers and students alike. Academic writing continues to be a very challenging, demanding, and yet a highly needed process in academia and beyond [26]. The degree of difficulty of academic writing is correlated with the degree of confusion about it. Students consider academic writing multifaceted and unnecessary, while teachers consider its marking laborious and non-rewarding [31]. Understandably, the process of providing formative feedback to student's guides teachers towards considering all the aspects of academic writing e.g., content and accuracy of language use and coherence or rhetorical effectiveness [29]. Nevertheless, in most cases, teachers make a self-consciousness decision to focus more in one of the aspects of writing than in the others, mainly to tackle one issue at a time so students can learn and grow. As a result, students develop those aspects of writing considered as important by teachers and get equipped with new sets of skills to be used to their benefit [18, 31].

Critical thinking is a crucial thinking skill used to describe the human ability to think about a problem and the possible solutions associated with it [21]. Moreover, critical thinking relates to the acceptance of opposing views and getting equipped with the required set of skills needed to participate in debates or to write academic essays, reports or thesis. Some of the processes that partake in the development of critical thinking are questioning, forming concepts, reasoning, solving problems and making decisions and judgements [28, 29, 32]. Critical thinking displays a relationship of informal apprenticeship between the substantial corpora of information present in the content, and linguistically related components of writing [48, 51]. The application of critical thinking in

academic writing through coherence and cohesion is a complex and multifaced process that involves a multitude of reader and text-based features [51]. Coherence is strongly correlated with the ability of the writer to choose a topic, read about it and finally write his thoughts in a way that is understandable and liked by others. A smooth flowing text enjoyable to read is the oasis of every academic writing teacher [14, 27, 36, 48].

Nevertheless, since summative marking of academic writing is highly subjective, we as teachers often discuss coherence in an ambiguous or inadequate manner [8, 18, 19]. Moreover, critical thinking is not only a very valuable skill to develop but it is also one required when completing the content of an assessment, and when paying attention to the accuracy of referencing. While it is a well-known fact that factors involved in the development of critical thinking are of nurture nature, it is substantially valuable to explore them with regards to the academic discourse [18]. Very often teachers focus on the mechanical aspects of academic writing when providing formative feedback to students, rather than in the content and appropriateness of rhetorical effectiveness of academic writing. Although language use is very important, development of thought through a diversity of ideas and extensive vocabulary creates the basis for a composition that attracts the attention of readers and allows the writer to express own thoughts appropriately [16, 21]. This creates a platform where the thought is appropriately developed, and when the application of it in academic writing reveals adequate levels of coherence.

2.2 Paraphrasing Tools and Plagiarism Instances

The dynamics of academic integrity, academic writing and the use of paraphrasing or translation tools are highly interconnected, therefore, it has become necessary to analyse them in correlation, aiming improvement of academic writing skills and subsequently reduction in plagiarism instances in HE. Multicultural students find academic writing quite challenging due to the constant increase in the amount of information that is being analyzed, its increased complexity, and commonly find a quick solution through the use of paraphrasing or translation tools to complete academic writing tasks [19, 20, 22]. Techniques of academic writing are a blend of sophisticated, discipline related vocabulary and complex and compound grammar focus; displayed in a formal and academic register [1]. Furthermore, the purpose of academic writing needs to be correlated with the ability to display academic knowledge in an academic format [14, 21]. As we live in an era when paraphrasing and translation tools are widely available, and when text-processing applications are easily found and accessed via the Internet, it becomes necessary to explore the use and the capacities of these tools to reproduce phrases, essays or reports and then 'spin', reprocess or translate [17].

The output from paraphrasing tools, article spinners and translation software can mislead people into thinking that these tools create a new form of original writing [43, 44, 48]. These forms of academic dishonestly not only damage the trustworthiness of the HE, but they also create dishonest employees and a corrupted society [16, 32]. Academic integrity in academia is a set of ethical and moral norms, which include maintaining standards and avoiding inappropriate behaviour such as cheating and e-cheating [25, 30]. It has been defined as a commitment made by students and teachers to uphold values such as honesty, trust, fairness, respect and responsibilities in the face of difficulties [4]. As a young industry, the HE in the UAE is yet to establish best

industry standards and practices across federal, semi-federal and private campuses in the UAE with regards to the use of paraphrasing or translation tools as well as academic writing and academic integrity. This is further challenged by the widespread of GenAI tools like ChatGPT and many new tools that are designed, improved and advanced continuously. Additionally, although different universities may have differing academic writing curricula or teaching and learning styles, it is believed that the core principles of the use of technology and of those of academic integrity apply to all students, as these principles are multi-dimensional in nature and need to be consistent across universities worldwide [19, 22]. Therefore, in order to create a solid foundation for good practices of academic writing and academic integrity, it is crucial to establish good academic writing practices and develop appropriate teaching and learning resources [21, 23, 25].

The use of technological innovations in the HE sector vary to a great degree as these tools are used by teachers and students alike for a number of functions and purposes. Educators use technological advances to facilitate the process of teaching, while students use them to facilitate the process of learning [18, 23]. The usage of technological advances may vary from providing students with feedback on their writing to editing writing their essays, or by utilizing the automated correction features which if not understood and explored appropriately could hamper a student's learning process rather than support it [13, 19, 24]. Academic integrity has always been a concern in HE and it has become an even more pressing issue with the extensive use of the Internet, and the prevalence of 'cut and paste' techniques and most recently the extensive use of GenAI tools [13, 25, 36]. Attempting to understand the perceptions of students about paraphrasing and translation tools, as well as their perceived difficulties in academic writing, serves the purpose of uniting the process of teaching and learning academic writing through the value of academic integrity. As the role of paraphrasing and translation tools is increased due to their widespread usage by students, it is important to explore the kind of difficulties that students face when producing academic writing tasks as well as the impact of paraphrasing or translation tools in students' ability to master academic writing skills in light of integrity. The plethora of paraphrasing, translation and GenAI tools used by students to complete written assignments is increasing, hence, it is important to understand reasons and ways why students opt to use these tools in academic writing [16, 24].

3 Methodology

For this study, the authors conducted a comprehensive literature review that covered the areas of GenAI gamification tools from its initial stages until the current year of this study. The approach was to formulate an understanding of the background information related to gamification and its use in HE. The authors focused on giving background information on the use of GenAI gamification tools in HE in a variety of contexts and purposes. The selection of the papers was conducted through Springer or SCOPUS indexed databases. Journal articles and conference proceedings were chosen based on their suitability and the purpose chosen by the researchers. The main purpose of choosing the articles was to provide an updated insight into the current uses of GenAI gamification tools in HE. There were a number of keywords and combinations of terms used to allocate the appropriate research sources. These keywords included GenAI, games, ESL games, higher education

and a variety of combinations of these terms. The first stage of sources collection was achieved through assembling a collection of articles that seemed relevant to the research questions. The second stage of sources analysis involved a careful analysis of the articles and a subsequent exclusion of the articles that lacked substantial discussion related to the content of this study. The focus was on analysing more recent papers; and the papers were categorised based on recurring themes that emerged while reading the articles thoroughly. Finally, as a result of this careful selection process, a total of 57 papers were included in this review and analysed in details for this study.

3.1 Sample

The target population for the study were undergraduate and post-graduate students studying in the United Arab Emirates and Turkiye. These students were at the time of the study enrolled in academic writing classes and were required to complete academic writing tasks as part of their university studies. All the students completed the survey anonymously and required around 20 min to complete both surveys. The surveys were conducted online. Students responded to Likert-scale items related to their difficulties and preferences in research and academic writing and the use of paraphrasing and translation tools. The survey was completed by 206 students. The researchers also explained to the students the purpose of each survey and informed students that they could withdraw at any time while completing the surveys. Students completed the surveys anonymously, and the incomplete responses were disqualified. Data were then extracted from the survey, organized and tabulated within the contingency tables using descriptive methods where necessary.

3.2 Data Collection

For this study, two surveys were used to collect data and answer the research questions. Survey 1 aimed to explore the needs of multicultural students to achieve the development of critical and analytical thinking skills when reading research sources and producing academic writing tasks. The survey questions were adapted from Razi (2015) and were comprised of 10 Likert-scale items was designed to collect the data aimed at analyzing critically how students perceived their existing needs related to academic writing, paraphrasing tools and plagiarism. The options given in the survey were: 1-strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, and 5 = strongly agree. The number of the respondents was 206 and students were enrolled in two international universities; 1) Turkiye and 2) United Arab Emirates. Data was collected in a period of three months through Qualtrics and the respondents were from over 60 nationalities. The survey questions were adapted from Razi (2015). The study utilized quantitative methods to obtain data from participants. The present study contributes to the ongoing research on measuring undergraduate students' attitudes, experiences and challenges faced when producing pieces of academic writing and the expectations of receiving formative feedback.

Survey 2 aimed at exploring concepts of development of academic writing skills and the tendency of students to use technology to conduct plagiarism. The survey was based on the Likert scale to critically analyse how students respond to their existing needs when it comes to the use of technology to complete academic writing tasks. The survey was

comprised of 8 Likert scale questions and the options given were: 1-strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, and 5 = strongly agree. For clarity of responses, both surveys were first piloted on a random group of students to test its validity and reliability, and as a result, some items were amended to make it easier for the multicultural students to understand. The purpose of this study was to explore the utilization of paraphrasing and translating tools used by multicultural students in academic writing.

4 Results

Table 1. Analytical and critical writing skills

No	Survey Item	1	2	3	4	5
1	Reading journal articles helps me understand another writer's viewpoint	10%	35%	23%	10%	32%
2	Reading journal articles helps me improve my academic vocabulary	5%	8%	12%	28%	47%
3	I understand written work better if I divide long paragraphs into shorter ones	12%	13%	14%	31%	30%
4	I understand better written work if I analyse one concept at a time	12%	28%	15%	24%	22%
5	I write in short paragraphs so others can understand clearly my viewpoint	35%	30%	10%	15%	10%
6	I prefer to analyse one concept at a time so my thoughts don't get confused	15%	12%	33%	15%	25%
7	My academic writing needs further improvements	15%	23%	25%	15%	37%
8	I can improve my academic writing skills by reading more	36%	20%	10%	23%	19%
9	I can improve my academic writing skills by writing more	20%	18%	22%	20%	20%
10	Receiving formative feedback from my teacher helps me my paragraph structure	25%	15%	10%	20%	30%

Table 1 shows the 10 Likert scale items and the percentage of each response. The options for the Likert-scale items were: 1-strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, and 5 = strongly agree. Questions were aimed at understanding the perceptions that multicultural students have with regards to the ways of developing academic reading and academic writing skills. Interestingly nearly 33% of respondents strongly agreed that reading journal articles supports their understanding of authors viewpoints, while exactly 48% of the students responded that their academic vocabulary is developed when reading journal articles. 30% of students strongly agreed that they performed better in reading when paragraphs were shorter, and that they understood the concepts better when they were analysed one at a time.

These findings give an indication that students lacked vocabulary and comprehension skills. Also, it is beneficial to acknowledge that students are aware of the need to develop their academic writing skills, as 37% of them strongly agreed that their academic writing skills required further improvements. However, interestingly the number of students who have faith in developing their academic writing skills by reading more was the same as the number of students who believed the opposite, which may indicate that more is required to raise awareness of the relationship between academic reading and academic writing skills.

Table 2. Academic writing, Technology and Plagiarism

No	Survey Item	1	2	3	4	5
1	Do you think academic writing is important for your university studies?	32%	15%	23%	10%	20%
2	Do you think academic writing is important for your individual self-satisfaction?	47%	28%	12%	8%	5%
3	Would you change an essay that you find online by using a text-processing software like "Grammarly" and present it as your own?	12%	13%	14%	31%	30%
4	Would you collect online essays with the same topic, paraphrase and summarize them and present them as your own?	22%	24%	15%	28%	12%
5	Have you ever used a text-processing e.g., 'Grammarly' to reprocess your essay, make it look more academic and then submit it as you own?	13%	30%	15%	12%	33%
6	Would you use 'Google translate" to translate big chunks of academic reading that you could not understand?	15%	10%	35%	13%	27%
7	Are you generally stressed before submitting a written assignment online?	38%	24%	25%	17%	15%
8	Do you think that a person who plagiarizes in university is prone to dishonesty in business world and social life?	35%	23%	13%	22%	17%

Table 2 shows the 8 Likert scale items and the percentage of each response. The options for the Likert-scale items were: 1-strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, and 5 = strongly agree. Questions were aimed at understanding the perceptions that multicultural students have with regards to the production of academic writing tasks, the use of technology to produce academic writing tasks and plagiarism. In Table 2 is seen that respondents did not view academic writing important for their individual self-satisfaction. The striking difference between the two extreme angles of the spectrum present in this question, with one being strongly agree standing at 5% and the opposing one being strongly disagree showing a value of 47%, demonstrates an urgent need to implement measures of improving the perceptions that students have about the value of academic writing in relation to their individual self-satisfaction.

Questions number three and number four were aimed at assessing students understanding of the use of paraphrasing and translation tools to complete academic writing tasks. Although a high number of students, 30% self-reported their intention of changing an essay by using 'grammarly', the answers that they gave for collecting essays from online sources and presenting them as their own was much lower and measured only 12%. Although this is a positive outcome as it displays students' lack of desire to cheat, yet the results pose another challenge in terms of academic writing skills and the use of paraphrasing and translation tools. Collecting a number of essays from online sources could to a certain extent mean that the respondents were willing to conduct some sort of research as they were going to include specific key words to collect essays that could sound as cohesive. Therefore, such exercise could possibly be used when trying to break down the concepts of searching for research articles online. Question number five and question number six were placed one after the other on purpose as to help the researchers generate an understanding of the ways that paraphrasing and translation tools are used by students to produce academic writing. Over 30% of students self-reported that they had used 'Grammarly' to reprocess their essay and presented it as their own. However, the question on the use of google translate tool to translate big chunks of academic writing that otherwise could be difficult to understand, produced rather interesting findings with 35% of students being neutral about it and 27% of students strongly agreeing to it. Therefore, this finding may open the path towards improving the vocabulary levels and the complexity of grammatical forms taught in academic writing skills to ensure that students are equipped with the adequate linguistic abilities to trust their own skill set. Finally, the question number eight generated results that indicate students' perceptions about plagiarism as a notion, placing it under category of the non-threating behaviour traits with regards to professional work experience. Interestingly 37% of respondents believed that plagiarism is not related to dishonest behaviour later in life, which indicates that more needs to be done to raise awareness about the personality changes related to plagiarism phenomenon.

5 Discussion

The study revealed that although academic reading and academic writing are highly interrelated, they are not perceived as such by multicultural students. This finding agrees with findings by Irkinovich and Izatullaevna [28]. Furthermore, the study highlighted the need to support students understanding of linguistic and skill related abilities required to complete academic writing tasks, similar to the study by Davis [6]. Furthermore, this finding seems slightly different from the study by Farouqa and Hysaj [10, 11] that highlighted the need to use technological tools to increase students' engagement levels. Specific measures and procedures related to academic writing, the use of paraphrasing and translation tools and plagiarism instances should be taken into consideration to guarantee improvement in academic writing tasks production by using ethically appropriate means [13, 14]. Factors like time management and external pressures are however not considered as intentional factors despite the possibility of leading to instances of plagiarism, mainly because these factors are conditioned by uncontrollable circumstances [5]. Awareness about issues related to the tendency to plagiarize and the tendency to cheat in the business

world need to be highlighted to guarantee a successful learning and work experience [7, 8]. Study by Handforth and Taylor [15] highlighted the requirement that students have to receive formative feedback before receiving summative feedback. The current study found that students require support, hence, formative feedback could be used by educators to provide those building blocks that students need to write successfully in an academic setting [11, 12].

This study found that practical guidance from teachers can support students in their journey towards improving their academic writing and research skills. This agrees to a certain extent with studies by Hysaj, Freeman and Khan [22] and Hysaj et al. [24] that emphasized the need to consider teachers and students as participants of equal importance. Creation of a supportive learning environment goes a long way towards empowering students with a variety of skills related to critical and analytical thinking and the ways they perceive plagiarism and the use of translation and paraphrasing tools [18, 23, 28]. Moreover, this study found that that despite efforts put in place in the two universities involved in this study, to raise awareness about the ethical use of paraphrasing tools, students use the paraphrasing and translation tools to produce academic written text and are not aware of their ethical implications related to the possible breach of academic integrity [22]. Lack of students' preparedness of completing academic writing tasks may be an added factor of encouraging the extensive use of paraphrasing and translation tools. This was another challenge pointed out in the study by Eaton [8] and Hysaj [26].

As the current study mentioned the unawareness of the mechanics of academic writing and/or the process of conducting research are other factors that influence the decision of students to plagiarise [16, 19]. Critical and analytical thinking skills are demonstrated through comprehension skills and appropriateness of the adequate flow of thought when writing. The current study found that students lack comprehension skills, therefore, they opt to use the tools to paraphrase or translate big chunks of their academic writing texts, that subsequently constitutes plagiarism [7]. Therefore, these angles need to be considered when making amendments to existing curriculum and instructional design [21, 28, 29]. The participants of this study did not relate plagiarism with dishonesty, which agrees with the findings by Pecorari and Petric [41] and Peters [44] and opens the path towards adaptations in curriculum design and policy trying to build a bridge between concepts of trust and honesty on personal, academic and professional level. Finally, this study found that expectations of students and teaching staff are not in sync since students view plagiarism more as means to an end rather than a display of a dishonest behaviour and this agrees with the findings of Curtis [5] and Davis [6].

6 Final Remarks and Recommendations

Educators and policy makers in higher education agree that academic writing skills play a vital role in the academic success of students. This study highlighted the need to raise awareness about the ethical use of paraphrasing and translation tools, as well as recognize the extensive use of these tools. Furthermore, students are required to practice more academic reading and academic writing, as well as be aware of issued related to

the unethical use of paraphrasing and translation tools. Students are required to develop their academic writing skills and be aware of ways to develop their critical and analytical thinking skills.

It is valuable that students develop their linguistic abilities and increase awareness about the ethical considerations of the unethical use of paraphrasing and translation tools. Furthermore, students need to develop their individual academic writing skills. As this study found that students lack confidence when completing academic writing tasks, it is important to build students self-confidence during academic writing classes. This will not empower them with necessary linguistic abilities but it will also encourage them to be active participants of the teaching and learning process rather than distant spectators. Getting involved in the teaching and learning process may have a positive and last-longing effect on students' academic writing skills.

To better understand students' perceptions about their unique difficulties in academic writing, this study should be followed by another one that applies a qualitative approach e.g., focus groups or interviews. This will allow to measure the perceptions about the value of academic writing and its relation with students' academic satisfaction. Moreover, the curriculum design may include reflective tasks where students are expected to reflect on the process of academic writing as a whole e.g., finding articles, looking for key terms and conducting literature review. This can encourage the development of critical and analytical thinking influencing students personal and academic growth. Due to the time limitations, this study was focused on the quantitative data gathered from two questionnaires, another study needs to consider utilization of focus groups; which could provide insights into the ways of encouraging students to read academic texts aiming to improve the comprehension skills.

Another study needs to analyse the aspects of academic writing skills e.g., grammar, vocabulary, mechanics of writing that students may develop when using the paraphrasing and translation tools. Finally, this study highlighted that it is the responsibility of academic writing teachers to make pragmatic decisions when teaching reading and writing skills in higher education e.g., including more reading practice and adapting the curriculum to encourage students to read academic texts and subsequently improve their academic writing skills.

References

1. Babcock, R.D., Thonus, T.; Researching the Writing Center: Towards an Evidence-Based Practice, Revised Edition. Peter Lang International Academic Publishers (2018)
2. Bertram Gallant, T.: Part-time integrity? Contingent faculty and academic integrity. New Direct. Community Coll. **2018**(183), 45–54 (2018)
3. Bloom, B.: Taxonomy of educational objectives (1956)
4. Centre for Academic Integrity. The Fundamental Values of Academic Integrity, pp. 5–9 (2012). www.academicintegrity.org/icai/assets/FVProject.pdf. Accessed 15 March 2024
5. Curtis, G.J.: Guilt, shame and academic misconduct. J. Acad. Ethics **21**(4), 743–757 (2023)
6. Davis, M.: Publishing research as an EAP practitioner: opportunities and threats. J. Engl. Acad. Purp. **39**, 72–86 (2019)
7. Eaton, S.E., et al.: In memory of Tracey Bretag: a collection of tributes. Int. J. Educ. Integr. **16**, 1–20 (2020)

8. Eaton, S.E.: Academic Integrity in the Age of Artificial Intelligence (2023)
9. Farouqa, G., Hysaj, A.: Exploring faculty members perception of utilizing technology to enhance student engagement in the United Arab Emirates: technology and the ICAP modes of engagement. In: International Conference on Human-Computer Interaction, pp. 67–76. Springer International Publishing, Cham (2022)
10. Farouqa, G., Hysaj, A.: Active learning in the lenses of faculty: a qualitative study in universities in the United Arab Emirates. In: International Conference on Human-Computer Interaction, pp. 77–90. Springer International Publishing, Cham (2022)
11. Fendler, R.J., Yates, M.C., Godbey, J.M.: Observing and deterring social cheating on college exams. Int. J. Scholarsh. Teach. Learn. **12**(1), 4 (2018)
12. Fenton-Smith, B., Humphreys, P., Walkinshaw, I., Michael, R., Lobo, A.: Implementing a university-wide credit-bearing English language enhancement programme: issues emerging from practice. Stud. High. Educ. **42**(3), 463–479 (2017)
13. Hamam, D., Hysaj, A.: Technological pedagogical and content knowledge (TPACK): Higher education teachers' perspectives on the use of TPACK in online academic writing classes. In: HCI International 2021-Posters: 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part III 23, pp. 51–58. Springer International Publishing (2021)
14. Hamam, D., Hysaj, A.: The aftermath of COVID 19: future insights for teachers' professional development in higher education. J. Asia TEFL **19**(1), 303 (2022)
15. Handforth, R., Taylor, C.A.: Doing academic writing differently: a feminist bricolage. Gend. Educ. **28**(5), 627–643 (2016)
16. Hysaj, A., Elkhouly, A.: Why do students plagiarize? The case of multicultural students in an Australian University in the United Arab Emirates. In: ENAI Conference (2020)
17. Hysaj, A., Suleymanova, S.: The analysis of developing the application of critical thinking in oral and written discussions: the case of Emirati students in the United Arab Emirates. In: 2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE), pp. 819–824. IEEE (2020)
18. Hysaj, A.: COVID-19 pandemic and online teaching from the lenses of K-12 STEM teachers in Albania. In: 2021 IEEE International Conference on Engineering, Technology & Education (TALE), pp. 1–7. IEEE (2021)
19. Hysaj, A., Suleymanova, S.: Safeguarding academic integrity in crisis induced environment: a case study of Emirati engineering and IT students in a private university in the UAE. In: International Conference on Human-Computer Interaction, pp. 236–245. Springer International Publishing, Cham (2021)
20. Hysaj, A., Haroon, H.A.: Online formative assessment and feedback: a focus group discussion among language teachers. In: International Conference on Human-Computer Interaction, pp. 115–126. Springer International Publishing, Cham (2022)
21. Hysaj, A., Hamam, D.: Dimensions of formative feedback during the COVID-19 pandemic: evaluating the perceptions of undergraduates in multicultural EAP classrooms. In: International Conference on Human-Computer Interaction, pp. 103–114. Springer International Publishing, Cham (2022)
22. Hysaj, A., Freeman, M., Khan, Z.R.: Theory of planned behaviour in higher education: exploring the perceptions of multicultural ESL students about cheating. In: International Conference on Human-Computer Interaction, pp. 58–71. Springer Nature Switzerland, Cham (2023)
23. Hysaj, A., Hamam, D.: What does it take to develop critical thinking? The case of multicultural students in a digital learning platform. In: International Conference on Human-Computer Interaction, pp. 49–57. Springer Nature Switzerland, Cham (2023)
24. Hysaj, A., Farouqa, G., Khan, S.A., Hiasat, L.: Reducing stress through formative assessments: a case of the digital platform. In: International Conference on Human-Computer Interaction, pp. 486–500. Springer Nature Switzerland, Cham (2023)

25. Hysaj, A., Freeman, M., Khan, Z.R.: Teaching academic writing skills: a narrative literature review of unifying academic values through academic integrity. Concurr. Sessions **12**, 157 (2022)
26. Hysaj, A.: Exploring the impact of group video creation in multicultural students in the online platform. J. AsiaTEFL **20**(1), 127–131 (2023)
27. Hysaj, A., Freeman, M., Hamam, D.: Using AI tools to enhance academic writing and maintain academic integrity. In: International Conference on Human-Computer Interaction, pp. 57–66. Springer Nature Switzerland, Cham (2024)
28. Irkinovich, N.R., Izatullaevna, I.I.: Methods for teaching reading. J. Pedagog. Invent. Pract. **15**, 47–50 (2022)
29. Kiriakos, C.M., Tienari, J.: Academic writing as love. Manag. Learn. **49**(3), 263–277 (2018)
30. Khan, Z., Sivasubramaniam, S., Anand, P., Hysaj, A.: The role e-tools play in supporting teaching and assessments with integrity during the COVID-19 pandemic. In: European Conference on Academic Integrity and Plagiarism 2021: Book of Abstracts, pp. 52–54. European Network for Academic Integrity (ENAI) (2021)
31. Khan, Z.R., Balasubramanian, S., Hysaj, A.: Using the IEPAR framework-a workshop to build a culture of integrity in higher education. In: 8th European Conference on Academic Integrity and Plagiarism (ECAIP), pp. 4–6 (2022)
32. Khan, Z.R., Hysaj, A., John, S.R., Khan, S.A.: Transitional module on academic integrity to help K-12 students in the UAE prepare for next stage of education. In: Academic Integrity: Broadening Practices, Technologies, and the Role of Students: Proceedings from the European Conference on Academic Integrity and Plagiarism 2021, pp. 263–287. Springer International Publishing, Cham (2023)
33. Khalmetski, K., Sliwka, D.: Disguising lies – image concerns and partial lying in cheating games. Am. Econ. J. Microecon. **11**(4), 79–110 (2019)
34. Klimova, B.F.: Developing thinking skills in the course of academic writing. Procedia Soc. Behav. Sci. **93**, 508–511 (2013)
35. Lim, F.V., Phua, J.: Teaching writing with language feedback technology. Comput. Compos. **54**, 102518 (2019)
36. Macfarlane, B., Zhang, J., Pun, A.: Academic integrity: a review of the literature. Stud. High. Educ. **39**(2), 339–358 (2014)
37. Maki, P.L.: Assessing for Learning: Building a Sustainable Commitment Across the Institution. Routledge (2023)
38. McKinley, J.: Critical argument and writer identity: social constructivism as a theoretical framework for EFL academic writing. Crit. Inq. Lang. Stud. **12**(3), 184–207 (2015)
39. McPeck, J.E.: Critical Thinking and Education. Routledge (2016)
40. Moore, E.: Accuracy of referencing and patterns of plagiarism in electronically published theses. Int. J. Educ. Integr. **10**(1), 42–55 (2014)
41. Pecorari, D., Petric, B.: Plagiarism in second-language writing. Lang. Teach. **47**(3), 269–302 (2014)
42. Peters, M.A.: Academic integrity: an interview with Tracey Bretag. Educ. Philos. Theory **51**(8), 751–756 (2019)
43. Peters, M.A., Besley, T., Arndt, S.: Experimenting with academic subjectivity: collective writing, peer production and collective intelligence. Open Rev. Educ. Res. **6**(1), 26–40 (2019)
44. Prentice, F.M., Kinden, C.E.: Paraphrasing tools, language translation tools and plagiarism: an exploratory study. Int. J. Educ. Integr. **14**(1), 11 (2018)
45. Razi, S.: Development of a rubric to assess academic writing incorporating plagiarism detectors. SAGE Open **5**(2), n2 (2015)
46. Roberts, J.: Personalised learning in developing countries – is higher education ready? In: Electronic Conference Proceedings, 10th European Distance and E-Learning Network (EDEN) Research Workshop (2018)

47. Roe, J., Perkins, M.: What are automated paraphrasing tools and how do we address them? A review of a growing threat to academic integrity. Int. J. Educ. Integr. **18**(1), 15 (2022)
48. Rogerson, A.M., McCarthy, G.: Using Internet based paraphrasing tools: Original work, patchwriting or facilitated plagiarism? Int. J. Educ. Integr. **13**(1), 2 (2017)
49. Rogerson, A.M.: The use and misuse of online paraphrasing, editing and translation software. In: A Research Agenda for Academic Integrity, pp. 163–174. Edward Elgar Publishing (2020)
50. Singh, M.K.M.: International EFL/ESL master students' adaptation strategies for academic writing practices at tertiary level. J. Int. Stud. **7**(3), 620–643 (2017)
51. Strobl, C., et al.: Digital support for academic writing: a review of technologies and pedagogies. Comput. Educ. **131**, 33–48 (2019)
52. Strain-Moritz, T.E.: Perceptions of Technology Use and its Effects on Student Writing (2016)
53. Stoesz, B.M.: Lies, lies, lies: detecting deception and implications for investigations of academic cheating. In: Academic Integrity in the Social Sciences: Perspectives on Pedagogy and Practice, pp. 53–68. Springer International Publishing, Cham (2023)
54. Swan, M.: UAE Among Emerging Education Hubs (2014). http://www.thenational.ae/uae/education/uae-among-emerging-education-hubs. Accessed 17 September 2024

"Do As Your Teacher Tells You!" How Is AI Use Regulated in Nordic Higher Education Institutions?

Sigurbjörg Jóhannesdóttir[7(✉)], Päivi Kousa[8], Salme Korkala[8], Jenny Tarvainen[8], Kristín Björnsdóttir[7], Judit Háhn[8], Susanne Dau[5], Eva Irene Brooks[1,2], Jeanette Sjöberg[2], Skúlína Hlíf Kjartansdóttir[7], Nataliya Berbyuk Lindström[6], Sylvana Sofkova Hashemi[2,6], Mona Lundin[2], Shola Oyedeji[3], and Robin Støckert[4]

[1] Aalborg University, Fredrik Bajers Vej 7K, 9220 Aalborg East, Denmark
[2] Halmstad University, Kristian IV:S väg 3, 301 18 Halmstad, Sweden
[3] Lappeenranta University of Technology (LUT), Yliopistonkatu 34, 53850 Lappeenranta, Finland
[4] Norwegian University of Science and Technology, Trondheim, Norway
[5] University College of Northern Denmark, Selma Lagerløfs Vej 2, 9220 Aalborg, Denmark
[6] University of Gothenburg, Universitetsplatsen 1, 405 30 Gothenburg, Sweden
[7] University of Iceland, Sæmundargötu 2, 102 Reykjavík, Iceland
sigurbjorg@hi.is
[8] University of Jyväskylä, Seminaarinkatu 15, 40014 Jyväskylä Yliopisto, Finland
salme.korkala@jyu.fi

Abstract. This paper explores the AI policies, frameworks, and guidelines established by Nordic universities (from Denmark, Finland, Iceland, Norway, and Sweden). The Nordic countries share characteristics such as a strong infrastructure supporting educational technology integration. The aim is to find similarities, differences, and gaps between the guidelines of Nordic universities in order to make the guidelines more practical and valuable for teachers and students. The theoretical framework focuses on guidelines and principles based on AI ethics, especially in higher education.

A content analysis was conducted on 39 documents from 12 universities. The analysis reveals two major themes in the documents: (1) Use of AI in Higher Education and (2) Ethical principles of AI in Higher Education guidelines.

Nordic universities are committed to leveraging AI responsibly while maintaining ethical, pedagogical, and institutional integrity. The results of this study give valuable insights into what kind of issues should be considered in guidelines of ethical AI in higher education and how, in addition to principles, practical examples are also needed for both teaching and learning. The results show that more research, common discussion, and collaboration are required. Additionally, the absence of societal and environmental well-being factors indicates room for continuous development of the policies.

Keywords: Artificial Intelligence · Higher Education · Nordic countries · Artificial Intelligence Policies · Artificial Intelligence Ethics

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 194–218, 2025.
https://doi.org/10.1007/978-3-031-93539-8_14

1 Introduction

Generative AI is profoundly transforming higher education, impacting teaching methods, student learning experiences, administrative operations, and ethical considerations of academia. These changes prompt critical questions for higher education institutions (HEIs), such as how to support teachers and students in acquiring AI skills, how to adapt assessment methods, and AI's role in learning support and its potential misuse. Addressing these issues through effective regulation is challenging. Key regulatory concerns include deciding which aspects of AI use should be controlled, identifying who should set these rules—government agencies, university administrators, legal experts, AI specialists, educators, or students—and defining the regulation goals. Should regulations aim to balance understanding AI's capabilities and risks, uphold academic integrity, or encourage innovation? These questions also intersect with broader research ethics issues.

This study aims to analyze the regulatory frameworks governing generative AI use in higher education across five Nordic countries. By examining policy variations and commonalities among HEIs in Denmark (DK), Finland (FI), Iceland (IS), Norway (NO), and Sweden (SWE), we seek to uncover unique approaches and shared principles that shape AI governance in these educational contexts. The Nordic region, known for its commitment to inclusive education and ethical technology use, may provide valuable perspectives on AI governance frameworks that align with these values. We focus on regulations aimed at students and teachers, investigating the guidelines governing AI use in academic study and teaching contexts. Through an analysis of regulatory strategies within Nordic HEIs, this study aims to identify common themes, unique implementations, and gaps in these policies and, therefore, to illuminate the best practices and lessons for effective AI governance in higher education.

The second aim of this study is to explore and understand the approaches and strategies adopted by Nordic governments in regulating the use of artificial intelligence within higher education institutions. The goal is reached by identifying common regulatory themes, unique national differences, and the implications of these regulations for academic policy, technology integration, and educational outcomes. We also seek to compare how different Nordic countries tackle similar AI topics within their educational institutions, highlighting any country-specific approaches or solutions.

Artificial intelligence (AI) is a multi-disciplinary field within computer science focused on creating systems that can perform tasks typically requiring human intelligence. It encompasses various methodologies, including machine learning and deep learning, enabling machines to learn from data, recognize patterns, and make decisions [1, 2]. AI is characterized by the ability to understand, learn, and apply knowledge across a wide range of tasks, exhibiting human-like cognitive abilities [3]. These systems can perform tasks typically requiring human cognitive abilities, such as reasoning, learning, problem-solving, perception, and decision-making [3, 4]. For this paper, AI might be defined as computer systems designed to interact with the world through capabilities typically considered human [5, 6]. Today's manifestation of artificial intelligence in HE usually comes in various generative AI applications. The umbrella term generative AI refers to deep-learning models, a range of technologies, of which ChatGPT is an example, which can generate text, images, and other digitized media [7].

Various authors have contributed to the criticism of the use of AI in education. For instance, educators are encouraged to become more involved with AI-related decision-making since the issues of inequality and injustice have already arisen concerning AI technologies [8]. Also, the issues related to large language model (LLM) based AI tools, especially unreliability with facts [9] and challenges with meaning-making [10] due to AI's technical limitations, have been noted in previous research. On the other hand, plenty of research aims to find sustainable practices for AI use in education [11, 12]. The study is a continuum of research on AI regulations and practices in higher education.

2 Artificial Intelligence in Higher Education

2.1 The Challenges of AI in Higher Education

Integrating AI in higher education offers personalized learning experiences, enhanced student engagement, and efficient administrative tasks [13]. However, concerns exist regarding the potential undermining of the learning process and compromise of academic integrity due to the widespread adoption of AI in education [14]. For example, since OpenAI launched ChatGPT in November 2022, it has enormously influenced higher education teaching and learning by introducing innovative approaches while presenting risks [15]. The teachers are also concerned about the fairness and transparency of AI-driven assessments, the need to align AI applications with educational objectives, and the potential for AI to replace human educators [13]. In addition, the application of AI in education is poised to transform teaching methodologies [16] radically.

Furthermore, the technological challenges in implementing AI policies and frameworks in teaching and learning environments by universities include concerns about academic integrity, misleading and low-quality AI outputs, educational automation, data privacy, and the negative impacts on fundamental issues surrounding human rights and educational equity [17]. Technological challenges, including technical and integration challenges, ethical and fairness concerns, and misalignment with educational goals, also play a significant role in implementing AI policies and frameworks in teaching and learning environments [18]. In summary, the growing dependence on AI systems introduces fresh challenges, especially concerning transparency and trust [19]. AI remains an enigmatic and unpredictable technology for many, eliciting fears stemming from its unfamiliarity [16].

2.2 The Ethical Guidelines and Principles of AI in Education

The critical reflection on how AI should be developed and used ethically has continued almost as long as AI has existed [20]. For example, the existing research of Raina, Mundra, Prasad, and Mishra [21] highlights the need for guidelines that address user-centered concerns and the ethics of learning analytics and algorithms in AI in education [22]. Over the years, countless recommendations, guidelines, and principles have appeared internationally and nationally [23–26]. Although their purpose is to make the ethical use of AI easier while minimizing its risks and disadvantages, many are ambiguous and overlapping [27]. For that reason, most companies that develop AI-based tools,

for example, for education, have their ethical guidelines based more on listing challenges than on how ethical problems could be solved in practice [27, 28]. According to Waelen [29], the principled approach to explaining AI ethics has been criticized, for example, because it is mainly based on Western values. Additionally, the ethical principles often lack context, are too general, or do not explain how to implement them in practice [30]. Practical constraints to the successful integration of AI in education include technical and integration challenges, training and support issues, ethical and fairness concerns, cost and accessibility, transparency, privacy issues, and misalignment with educational goals [31].

The regulatory environment in the field of AI in education is becoming increasingly dynamic, with the adaptation of existing or the adoption of new comprehensive laws and legal frameworks at both the national and international levels to ensure the ethical, non-discriminatory, sustainable, and verifiable use of AI in education [32]. The European Commission High-Level Expert Group on Artificial Intelligence [24] has listed seven key requirements for trustworthy AI. These principles, which include individual and societal aspects, are used in this paper to support the analysis of the results.

1. Human agency and oversight
2. Technical robustness and safety
3. Privacy and data governance
4. Transparency
5. Diversity, non-discrimination and fairness
6. Societal and environmental well-being
7. Accountability

Integrating AI in higher education aims to revolutionize teaching, learning, and administrative procedures, leading to enhanced educational results. Ethical considerations in developing AI policies and frameworks for teaching and learning environments include ensuring responsible and safe deployment of AI systems, understanding the ethical implications of AI systems, and developing strategies to ensure responsible and safe deployment of AI systems [33]. The study emphasizes the necessity of implementing specific methods and policies to address the technological gap while supporting inclusive education strengthened by artificial intelligence, ensuring that all institutions can effectively deal with the difficulties associated with integrating AI to promote educational fairness and progress [34].

2.3 Nordic Context and Research Question

Higher education across the Nordic countries shares characteristics such as a strong infrastructure supporting educational technology integration (EdTech) due to their high student-to-computer ratios and excellent access to bandwidth and connectivity. This infrastructure supports integrating various technologies, including artificial intelligence, in educational settings. The use of EdTech in the Nordic region spans from preschool to higher education, indicating a broad acceptance and implementation of these technologies [35].

AI is increasingly being integrated into higher education globally, including in the Nordic countries. It can potentially transform educational methods and enhance administrative efficiency profoundly, but it also raises notable ethical and practical issues,

including data privacy and algorithmic prejudice [36]. While AI offers numerous benefits, addressing the challenges and ensuring its responsible use is essential to maximize its positive impact on higher education.

Gordienko and Bagraioni [37] revealed six main research trends in the context of higher education and AI. Three of the areas are the development of regulations for the use of AI, improving operational processes at universities, and enhancing the literacy of teachers and students in AI. Universities are amending and shaping policies regarding AI implementation, with faculty being able to determine how AI tools are used in their courses. These trends could encourage innovation and creativity but may lead to inconsistency in students' learning experiences due to a lack of standardization [18]. More research is needed regarding the ethical use of AI to provide clear policies and resources for universities, as institutions vary significantly in terms of faculty guidance [38]. Kim and Wu [39] found that challenges faced in developing and enforcing AI policies include the lack of consistency in existing policy practices. Especially in the light of emerging ethical considerations, this calls for urgent, multi-disciplinary policy development to harness AI's potential while mitigating its risks.

In Nordic countries, there has been a significant increase in the utilization of artificial intelligence among students. In Norway, for instance, the proportion of students who reported using AI in their academic work rose from 61% in 2023 to 81% in 2024. Despite the substantial rise in AI adoption, 67% of students expressed that they had received inadequate training in the practical application of AI tools. Notably, those who had undergone AI training demonstrated a greater propensity to integrate AI into their academic pursuits, and the extent of AI utilization varies across disciplines and educational programs [40].

These findings emphasize the imperative for educational institutions to enhance training and support mechanisms, ensuring that students are adequately prepared to navigate the evolving academic and professional landscapes shaped by AI advancements. These technological advancements inevitably introduce multiple risks and challenges, progressing faster than policy discussions and regulatory frameworks can keep up with [5]. Our study conducts a content analysis of AI policies and guidelines in Nordic higher education institutions to identify best practices that strengthen AI governance and promote the responsible and ethical use of AI technologies within higher education. The findings aim to inform and support universities in refining their AI policies and guidelines, ensuring they remain effective and aligned with evolving technological and regulatory landscapes. This paper explores the following question: *What recurring themes emerge in the AI policies, frameworks, and guidelines established by Nordic universities?*

3 Data and Background

This section introduces the selected countries' digital contexts alongside the documents and methodology used in the study. The chapter outlines the data collection and analysis processes, highlighting the key sources and employed techniques. Additionally, it briefly discusses the ethical and methodological considerations that shaped the research.

3.1 Denmark

The Danish population and society are among the world's most digitalized and have the second most digitalized public sector [41]. In 2024, the Government launched a strategic initiative based on three guiding principles for artificial intelligence, together with four concrete initiatives aiming at speeding up the responsible use and development of artificial intelligence in Denmark [42].

The vision emphasizes that artificial intelligence development and use must always occur with citizens in focus. Further, Danish companies must be globally competitive, and Denmark should be a world leader in using artificial intelligence in the public sector. The initiatives include the establishment of a digital task force, a new center for artificial intelligence in society, a platform to boost the development of a sage, a transparent Danish language model for the use of companies and authorities, and the digitalization of Danish text data into an open-source solution. The governmental strategic initiative mentions relatively little of the educational area as part of the strategy, despite some of the universities' involvement in developing the platform and funding opportunities.

3.2 Finland

Regarding the level of digitalization, Finland has been among the first in the EU with its highly digitally skilled population, comprehensive digital policies, and well-functioning e-government services [43]. As a part of the EU, Finland is following the EU Act [44] when considering using AI. In Finland, there is no governmental legislation on AI. Finland is halfway through a two-year preparation process to make a legislative proposal for the national implementation of EU-based AI regulations [45].

There are 24 universities of applied sciences and 14 universities in Finland. The Finnish documents covered in this paper are from the University of Jyväskylä (JYU) and Lappeenranta-Lahti University of Technology (LUT). The University of Jyväskylä was selected because JYU's faculty, Jyväskylä School of Business and Economics (JSBE), was the first to publish AI regulations in Finland (dated Jan. 26, 2023) and was among the first ones in the EU. LUT is part of the study because it is one of the multi-campus universities in Finland, where there is much technology-mediated communication between teachers and students.

Each university has its own AI policy. Additionally, some faculties and units have created discipline-specific AI guidelines due to the differences in how the large language model-based tools affect students' learning outcomes, assessment, and evaluation in various disciplines. For instance, the general JYU guidelines state that AI is prohibited when language is assessed. There are situations when this is sensible, but language is constantly part of communication and language studies assessment. Nevertheless, AI can still be used mindfully, for example, explicitly introduced to support these skills, e.g., academic writing.

All the guidelines of JYU were created in 2023 except the additional guidelines for teachers, which were created by the Centre for Multilingual Studies (Movi) and finalized in August 2024. At the end of 2024, JYU had six guidelines, which have been updated regularly. The policy on the use of artificial intelligence of the University of Jyväskylä is the underlying document for everyone at the university. The other guidelines

provided by different units are based on this policy. JYU's instructions and guidelines for using AI-based applications in studies target all university students. When a faculty or another unit has no guidelines, these student guidelines are used predominantly. The units that have created more detailed guidelines for the students are Movi, JSBE, and the Department of Language and Communication Studies. The Department of Language and Communication Studies guidelines focus specifically on thesis writing. Only Movi has created targeted AI guidelines for the teachers. The purpose of these guidelines is to help teachers discuss and utilize AI in their courses. LUT University created an AI-based tools guideline document for the students and teachers to ensure the ethical and responsible use of AI tools in teaching and learning.

3.3 Iceland

Iceland has made significant strides in integrating digital technologies and expanding the internet. Nearly all the population (99% in 2020–2024) is online [40], compared to 70% on average worldwide [46]. The Government's focus on digital services strengthens society, boosts competitiveness, and fosters innovation [47]. Additionally, efforts emphasize public knowledge of digital tools, including AI.

AI has played a crucial role in integrating the archaic Icelandic language into technology, which is noteworthy in the case of a small nation with a population of 370,000 inhabitants. The Government's AI policy [48] balances innovation with social values, ensuring AI technologies benefit all citizens while safeguarding democratic principles. Priorities include education, ethical AI frameworks, and international collaboration. Iceland has partnered with global tech companies, including Google and Microsoft, and eventually partnered with OpenAI to enhance Icelandic capabilities in AI [49, 50].

The Ministry of Higher Education, Industry, and Innovation in Iceland [51] has spearheaded policy regarding language technology, which focuses on AI for public benefit, competitiveness in business, modernization of education, transformation of public administration, and AI in healthcare. In education, priorities include AI-driven learning tools, specialized educator training, and AI literacy programs. Higher education initiatives support MSc programs in AI and data science. These actions aim to enhance accessibility, modernize learning, and equip citizens with future-ready skills.

3.4 Norway

Norway currently does not have specific laws or regulations directly regulating AI. The legal framework is mainly technology-neutral and flexible, indirectly encompassing AI. As an EEA member, Norway is expected to implement the EU AI Act [52] eventually. However, it has some concerns about its relationship with other legal frameworks and its impact on SMEs, which was addressed through a Position Paper [53]. Further initiatives shaping Norway's approach toward implementing AI include signing the treaty by the Council of Europe's Framework Convention on AI [54] and partaking in the Nordic prime ministers' vision of having a unified and coherent approach toward developing robust, ethically sound AI solutions. This strategy is vital for aligning with EU regulations and setting a global benchmark in ethical AI [55].

In recent years, Norway has undergone an extensive digital transformation. In practice, we currently live in a thoroughly digitalized society. Norway has a well-developed communications infrastructure, and most people can access high-speed broadband. The OECD Digital Government Index [56] assesses the efforts made by governments to establish the foundations necessary for a digital transformation of the public sector. In 2023, Norway ranked fourth among OECD countries [57].

Norway aims to become the most digitalized country in the world by 2030, leveraging digitalization and AI to address societal challenges, enhance public services, and foster a competitive business sector. The national strategies for digitalization [58] and AI [59] emphasize digital competence, secure infrastructure, ethical AI use, and the integration of AI in higher education.

Several initiatives have been initiated to scaffold these strategies' implementation processes and enhance the collaboration within and between the private/public and university [60–62] sectors.

On a governmental level, the Norwegian Agency for Public and Financial Management [63] has created a marketplace for cloud services that will make it easier for businesses to use secure, legal, and cost-effective cloud services. They call this process the "Cloud journey" or skyreisen [64]. The cloud journey is a process that a business can use to prepare a cloud strategy and prepare, procure, and implement cloud services. (The process is divided into four phases with relevant guidance, templates, and articles.)

The Norwegian Government foresees designating a national supervisory authority per the EU AI Act proposal. It has not yet been decided which regulatory body will remain the primary regulator of AI, and several bodies have expressed interest. The Norwegian Data Protection Authority may be a potential candidate, as it provides a "sandbox" for privacy-friendly innovation and digitalization. In the spring of 2024, the sandbox was opened to introduce, develop, and use generative language models [65]. One of the accepted projects was the Microsoft 365 CoPilot project at the Norwegian University of Science and Technology (NTNU) [66]. The momentum and results from the project have generated a solid foundation for developing and implementing rules and regulations for AI for NTNU, which is expected to be released around May 2025.

3.5 Sweden

Sweden has a high internet penetration rate, with around 98% of the population having access to and 95% of the population using the internet [67]. This widespread access supports a robust digital infrastructure, enabling extensive use of online services and digital tools across various sectors, including education. Sweden has been proactive in integrating AI across different domains. The national AI strategy, launched in 2018, emphasizes education, research, innovation, and infrastructure. AI Sweden, a national center for applied AI, supports numerous projects and initiatives to foster AI adoption. In Sweden, there is a strong emphasis on digital literacy, with various initiatives to ensure that citizens can effectively use digital tools and services. In 2017, the Swedish Government adjusted curricula and syllabuses to strengthen digital literacy. These changes include introducing programming in various subjects, enhancing students' problem-solving skills using digital technology, and promoting the critical evaluation of digital

sources. However, despite high internet penetration, there are still disparities in access and digital literacy among different socio-economic groups and rural areas.

In Sweden, the regulations for AI in higher education are shaped by national and European frameworks. The Higher Education Act [68] and the Higher Education Ordinance [69] provide the general regulatory framework for higher education, including digital work and AI aspects. Effective March 2024, the EU AI Act sets stringent requirements for AI applications, especially those involving personal data and high-risk AI systems. It impacts how universities can use AI tools, ensuring compliance with privacy and safety standards. Sweden's national AI strategy from 2018 [60] emphasizes the integration of AI in education and training. It encourages universities to offer AI-related programs and ensures that technical and non-technical courses include AI components. Individual universities may have institutional policies and guidelines on using AI, focusing on academic integrity, data privacy, and ethical considerations. These regulations aim to balance the innovative use of AI with the protection of personal data and academic standards. Sweden has approximately 50 universities, university colleges, and individual education providers. Higher education and research in Sweden are primarily conducted at state universities and university colleges.

4 Data Collection and Analysis

This study uses qualitative content analysis to examine artificial intelligence (AI) policy documents from higher education institutions (HEIs) in Denmark, Finland, Iceland, Norway, and Sweden. Content analysis is a research method that systematically examines patterns within textual data. It provides an objective approach to analyzing large volumes of content, allowing researchers to identify underlying themes, trends, and variations [24]. The systematic analysis of policy documents and guidelines across national contexts highlights variations regarding higher education AI policies.

Data were collected from official AI policy and guidelines documents of 12 universities and national AI policy documents of five Nordic countries (Sweden, Norway, Denmark, Finland, and Iceland). The selection criteria included selecting documents that address AI policy or guidelines at the university level, coursework, assessment, ethical responsibilities, and national level.

The universities were selected based on access to and availability of policy documents and guidelines. Two universities per country were selected in Denmark, Finland, Norway, and Sweden, while four universities were chosen in Iceland. The average number of students in these universities is approximately 20,000, the largest institution with over 50,000 students and the smallest having less than 700 students. A total of 39 documents were analyzed.

Data analysis followed a qualitative content analysis approach with manual open coding, category grouping, and theme construction. This study applied an inductive approach where themes emerged from data rather than predefined categories. The first step was an open coding process, which was used to assign descriptive labels to the data. These activities were done by the first five authors, with subsequent refinements carried out iteratively by all authors to ensure consistency. Once open coding was complete, related codes were grouped into broader categories that shared common characteristics and refined iteratively by all authors not involved in the initial grouping. Finally,

themes were constructed from the categories to provide deeper insight into the research questions. This approach aided in a comparative analysis of AI policies across Nordic institutions to identify key trends and national variations in AI policies and guidelines for education among the participating universities.

A detailed examination of the guidelines highlighted the measures necessary to safeguard ethical standards and promote a responsible AI usage framework across academic settings. Firstly, six recurring themes were found through content analysis of the Nordic university guidelines. They were Data Privacy, Intellectual Property, Transparency and Accountability, Intentional Misuse, Unintentional Misuse, and Disciplinary Actions. However, there was an overlap between the thematic groups, which resulted in challenges to determine and place "intentional and unintentional misuse" and "disciplinary actions" into the correct ethical categories. Because there is much research and countless amount of guidelines for AI ethics, it was decided to do one more analysis according to the seven EU principles: 1) human agency and oversight, 2) technical robustness and safety, 3) privacy and data governance, 4) transparency, 5) diversity, non-discrimination and fairness, 6) societal and environmental well-being, and 7) accountability [70] to find out if there was more comprehensive information in the thematic groups. The second and third authors of this paper ran this analysis. By running an analysis based on the seven EU principles of AI ethics, it was also possible to find out how the thematic groups aligned with the EU principles and if there were any gaps between principles. This analysis revealed that all the thematic groups that were first found, including misuse and disciplinary actions, were mainly related to privacy and data governance (3) as well as transparency and accountability (4 and 7). Some of the principles overlapped, which made categorization challenging. Without the theory-driven content analysis, the theme of human agency and oversight (principle 1) would have been overlooked. Diversity, non-discrimination, fairness (5), and societal and environmental well-being (6) were left off after the content analysis phase because they were not mentioned in the guidelines.

4.1 Ethical and Methodological Issues

During the research process, multiple researchers involved in the project were conscious of the importance of analyzing the documents systematically and objectively. Systematic documentation during data analysis is vital to prevent researcher coding, categorization, and interpretation bias. Furthermore, multiple researchers strengthen the triangulation of findings since they focus on various aspects of the data, allowing for a more comprehensive and nuanced interpretation [24].

5 Findings – Recurring Themes

The findings in this section are drawn from a content analysis of 39 policies and guidelines. This analysis highlights how AI is governed in Nordic universities by uncovering patterns and themes in their policy documents and guidelines. Six key themes emerged, categorized into two areas. The first category addresses the use of AI in higher education, including (1) teaching and learning, (2) research and management, (3) implementation and training. The second category focuses on ethical principles, covering (4) human

agency and oversight, (5) data privacy, intellectual property, technical robustness and safety, and (6) transparency and accountability.

5.1 Use of Artificial Intelligence in Higher Education

Teaching and Learning. Higher educational institutions in Denmark, Finland, Iceland, Norway, and Sweden have created guidelines about using AI in teaching and learning to balance ethical and academic responsibilities. A common requirement across the documents analyzed is that teachers are responsible for informing students about the appropriate use of AI in their courses, and information about AI use should be included in course syllabi. furthermore, frameworks and guidelines emphasize individual teachers' authority to determine AI usage in their courses.

Most universities require teachers to define the boundaries of prohibited AI use and justify these restrictions. Teachers must understand AI's capabilities, limitations, and ethical implications to achieve this. For example, teachers at the two Swedish universities under study are expected to develop competence in ethical AI and data use. Icelandic teaching guidelines are less prescriptive and more reactive, reflecting the country's decentralized education system, potentially allowing for flexible AI use in teaching and learning in higher education. The University in Oslo emphasizes the importance of using its own developed and locally running version of ChatGPT as a tool to support teaching and learning. It encourages teachers to familiarize themselves with AI technology and its ethical, social, and legal implications before using it in their courses. Teachers are advised to use this local ChatGPT version for lesson planning, developing assignments, and creating learning activities.

Finnish guidelines mandate that all AI applications used in courses must be available and accessible to all students free of charge. Additionally, these guidelines highlight the impact of AI on evaluation, suggesting that teachers might need to re-evaluate assignments and learning outcomes in consideration of these rapidly developing technologies. While most universities stipulate explicit prohibitions against using AI in exams or careful considerations, some, for example, Aalborg University, allow AI like other aids in exams unless otherwise specified by the teacher. While it is the teachers' responsibility to inform students about permitted and prohibited AI use, students are responsible for staying informed about central and local guidelines regarding the use of generative AI.

The teaching policies and guidelines analyzed are embedded within broader academic integrity frameworks to prevent misuse, such as plagiarism or unauthorized AI-assisted work. The universities emphasize that AI should serve as a learning tool rather than a shortcut that undermines independent work and intellectual engagement. For instance, at JYU in Finland, AI use is encouraged and defined as a work-life skill students must acquire. However, they should only use AI as an aid rather than a replacement for academic work.

Across the Nordic countries, the AI guidelines state that students are ultimately responsible for how they engage with AI. For example, the statement "It is your responsibility!" in Aalborg University's guidelines carries a clear legal and ethical focus, emphasizing that students are accountable for using AI in their educational activities. The Nordic guidelines share that AI should not be used to replace one's thinking, reading,

and writing. For example, the JSBE guidelines in Finland state that "the purpose of university studies is to develop critical thinking skills, and this skill does not develop if thinking is outsourced to AI." The guidelines in Finland, Norway, and Sweden specifically emphasize the importance of developing critical thinking skills related to using AI, which is understood as a part of responsible expertise.

All countries stress the importance of transparency, requiring students to disclose AI-generated content when necessary. Presenting AI-generated work as one's own (misusing AI) may lead to academic consequences. The AI guidelines of Aalborg University emphasize that students need to be aware of the extent to which they are permitted to use generative AI services in their academic work. Three situational examples are listed in the document that students are advised to pay attention to, including the potential consequences of failing to attribute AI-generated contributions correctly. These examples help students to avoid ambiguity in what is allowed and what constitutes academic misconduct.

In some countries, AI is suggested to feed students' creativity and help them create new thoughts and ideas, considering that they need to develop those ideas further. Using AI as a sparring partner and inspiration is especially central in the Finnish and Norwegian AI policies. For instance, the University in Oslo (UiO) suggests that GPT UiO can help students by providing feedback on written work, generating ideas, and offering explanations. UiO stresses the importance of students being in control of their learning and using AI as a support tool: "Let AI work with you, not for you." The UiO guidelines recognize AI's potential for providing feedback on written texts and simplifying complex information. The guidelines also emphasize AI's role in supporting students with disabilities, like Autotekst, a UiO-developed tool for automatic transcription of speech to text, powered by Whisper from OpenAi.

Only a few universities acknowledge that students can use AI specifically for creative work. For example, in documents from Aalborg University, it is stated that AI can generate code and create images, videos, and audio if students adhere to current license terms and conditions. At the Iceland University of the Arts, students can use AI in creative work in consultation with their teachers. When they do, they are required to adhere to legal and ethical regulations. Swedish guidelines acknowledge that generative AI tools can produce creative works. However, such work should be managed in educational settings to ensure originality, compliance with intellectual property rights, and academic integrity.

There are differences between the universities where AI tools are allowed or recommended to support learning. In Denmark, only GDPR-approved digital tools, such as Microsoft CoPilot and Dall-E, are favored. ChatGPT and other generative AI are permissible in Iceland to support learning activities. There is a variety of approaches for Norwegian Universities. The University in Oslo allows the use of their own developed and adapted ChatGPT or Whisper from OpenAI and is building guidelines to support using these dedicated tools in teaching and learning. NTNU is working with the Government and running a Sandbox pilot test of Copilot and its many features, functions, and integrations. The Sandbox project has produced a comprehensive overview of the opportunities, challenges, and critical topics for AI implementation by publishing several reports and workshop results for the whole process. The overview can then be used

as a checklist to create guidelines for implementing and using AI for HE and public and private organizations.

Overall, the Nordic institutions have different policies and guidelines for using AI in teaching and student learning, but all emphasize academic integrity and responsible use of AI.

Research and Management. AI use across the twelve nordic universities is expected to align with sound scientific practices. Teachers and students are expected to investigate and verify their sources and whether the tools used in teaching and research utilize generative AI-based enhancements.

The Danish universities and two Icelandic universities do not elaborate on research-related aspects of AI use in their guidelines. However, most other Nordic universities have done so primarily regarding graduate students' research projects. For example, at LUT, students can use AI for brainstorming and language refinement or editing. However, LUT does not recognize the use of AI as a reliable scientific source for students. Finnish AI guidelines allow students to use AI during the pre-writing phase of their thesis. It can help students generate research and interview questions, define topics, and explore related concepts. It can support text quality in writing through proofreading, editing, translation, and refining research materials such as data examples, research instruments, and consent forms. Students should not use AI tools to write the MA thesis abstract because that is their maturity exam. In addition, students must acknowledge the use of AI in a statement attached to the thesis.

Two Icelandic universities stipulate in their guidelines that AI can be used for research purposes. For example, Iceland University of the Arts allows AI to be used as support during the initial stages of the research process and for proofreading documents. The University in Oslo calls for responsible AI use in research, particularly regarding compliance with data protection laws and transparency. Furthermore, Norwegian universities host various research centers like the Norwegian Open AI Lab and the Norwegian Research Center for AI Innovation.

One Swedish university explicitly stipulates that supervisors of PhD students are responsible for providing access to necessary AI tools. In terms of assessment, the University of Iceland employs a rubric to evaluate how students use AI in research tasks such as data collection, analysis, argumentation, writing, and editing. This rubric categorizes projects into three levels based on management: 1) managed entirely by students, 2) managed jointly by students and AI, and 3) managed entirely by AI.

Most of the documents emphasized teaching, learning, and ethical practices. Less focus was placed on administrative management. For example, only one university in Iceland includes administrative management in its guidelines regarding AI usage. Bifröst University outlines that staff can use ChatGPT and other similar applications to draft emails, reports, and plans, as well as communication and collaboration between departments. All current ethical and academic rules and guidelines also apply to AI.

Swedish guidelines align the management of AI with previous guidelines for the use of IT tools at the university level, for example, issues such as access to AI tools for students, the use of AI systems regarding sensitive personal data, the dynamics of the guidelines since AI technology is changing rapidly, and ethical and legal requirements (e.g., AI must comply with the AI Regulation, the Data Protection Regulation, and

other relevant legislation). Similar issues are raised in Finnish guidelines, which declare responsibility for procuring AI tools for universities and updating the guidelines.

The absence of direct references to management in most of the documents suggests that the use of AI in management practices at these institutions has not been distinctly identified and addressed.

Implementation and Training. Implementing AI training in universities varies within and across the nordic countries, with some institutions offering structured courses and others focusing on general guidance. While Norway and Denmark provide structured training approaches, sweden and Iceland emphasize teacher-led learning. finland stands out for its early AI literacy course, and denmark is actively piloting new training programs. A common theme across all institutions is recognising AI's impact on education and the need for continuous training and ethical awareness to ensure responsible and effective AI use in academic settings.

In Norway, the University in Oslo follows a structured approach to AI implementation based on its "own" AI tools. Hereby, they offer recommendations for implementing AI, including defining target groups and their needs, establishing a "prompt school" to teach users how to use AI, and conducting thorough risk and privacy impact assessments. NTNU, the other Norwegian university in the present study, provides practical guidance on AI tools and fosters engagement through webinars and forums but lacks specific training programs. Both institutions highlight the importance of ongoing education and ethical considerations related to AI technologies.

Sweden does not provide a specific list of implementation steps but outlines key principles emphasizing clear communication and training. Training programs are implied rather than explicitly stated, with guidelines suggesting that students and teachers need support to understand and utilize generative AI effectively. The focus is on informing and training educators and learners while ensuring transparent communication of expectations and rules for AI use.

In Iceland, Bifröst University stands out as the only institution mentioning in its documents that it offers formal AI training. In contrast, other universities leave the responsibility of AI education to individual teachers. Although the analyzed documents provide minimal training information, all universities and their institutions have hosted AI information and training events in recent years.

In Finland, JYU offers training for the staff and the students. Since the teachers are obligated to support the students' AI use, it is important to offer support for the teachers. These activities are organized by offering knowledge and training on AI in teaching, especially for the teachers. Both online material and a course are offered. The students can explore AI literacy and its applications in their academic and professional lives in a course focusing on those issues. This course was one of the first AI literacy courses available in Europe. The students also get guidance from their teachers during courses, especially during communication and language studies provided by Movi.

Denmark has recently initiated pilot training programs at the University College of Northern Denmark (UCN), where teachers have access to skills development and guidance on approved generative AI tools. Students receive an internal online course on generative AI, focusing on prompt engineering. Aalborg University offers a comprehensive online course accessible to both teachers and students. The course is divided into

two sections: the first part introduces fundamental AI concepts such as Large Language Models (LLMs) and Reinforcement Learning from Human Feedback (RLHF), while the second part emphasizes practical applications of AI, particularly in text generation, tutoring, and content creation. Ethical considerations such as bias, misinformation, and over-reliance on AI are acknowledged but not deeply explored.

5.2 Ethical Principles of AI in University Guidelines

In the realm of artificial intelligence within academic environments, the ethical management of data and the potential for misuse stand as pivotal concerns. As AI continues to integrate into the fabric of academic life, understanding these dynamics is crucial for maintaining the integrity and trust essential to educational institutions. This chapter delves into the complexities of ethical and data considerations, exploring how ethical principles appear in the guidelines and what can be concluded from them, for example, regarding intentional or unintentional misuse. The predominant theme across the policy and guideline documents is reinforcing existing ethical standards in learning, teaching, and research.

Human Agency and Oversight. All the examined HEI policy and guideline documents are user-centered because they are developed to support students and teachers in AI's ethical use and avoid academic fraud. even the names of the documents stress that they cover issues of how students, teachers, and other academic experts should be using AI tools. Thus, they all emphasize the user's responsibilities when using AI tools and indicate that students and teachers play a crucial role in ensuring AI's safe and ethical use. Because teachers and students have this central role, and their actions define if the AI tools are used safely and ethically, it is evident that the human agency and oversight category is intertwined with other categories in the results of this research.

Another similarity is that all the documents expect students and teachers to be aware of the limitations and possibilities of AI tools usage, and this way, they expect the AI users to have the capability to understand how, for instance, AI tools based on large language models work. The third similarity is that all five countries' documents state that students and teachers are expected to follow GDPR when using AI tools in their studies or teaching. When the guidelines focus on data protection laws, they help AI users actively prevent or reduce risks in academic contexts. On the other hand, AI users are also seen as responsible for understanding the restrictions data protection laws bring to data use.

The difference between the documents is how they present teachers' role in advising students on using AI. In Norway, the guidelines of the University in Oslo stress teachers' agency by expecting teachers to ensure that students understand the privacy implications of using AI tools and avoid including sensitive personal information in their prompts. Other guidelines do not have this kind of teacher responsibility. Instead, they pinpoint that students are responsible for their learning tasks during their studies.

The Finnish HEIs have more detailed guidelines and policies than other countries in this study. They give exact directions on what students and teachers must and must not do, for example, when writing a thesis. A Finnish guideline (JYU Movi guideline for students) and a Danish guideline (Aalborg University) bring up student responsibility,

especially in the context of group work, and the Finnish guideline also in the context of peer feedback. Both group work and peer feedback are social studying contexts where the students need to have a shared understanding of how and why they can use AI in their studying tasks.

All the guidelines allow students, teachers, and researchers to use AI tools if the use is reported and ethical principles are followed. The Finnish guidelines also present AI literacy and the ability to use AI tools effectively as a work-life skill every student should master. The LUT guidelines even express that the students should be oriented into the matter throughout their studies.

One of the themes connected with human agency and oversight is critical thinking. The theme is stressed in the guidelines of Finland, Norway, and Sweden but not so profoundly in the documents of Denmark or Iceland. The Finnish guidelines point out that using AI tools requires critical thinking skills, and by using AI, students should not outsource learning to AI. A Norwegian guideline (NTNU) summarises this by advising students to assess and verify AI-generated information critically and explain why and how they use AI. According to this guideline, this helps graders assess the students' methodological approach. This example also indicates that the students' and teachers' actions affect each other when considering the use of AI tools.

All in all, when contemplating the aspect of human agency and oversight, the guidelines from all the Nordic countries seem to view AI as a support tool in the learning and writing process. They also see students and teachers cooperating when learning to use AI tools ethically in their academic processes; after all, many are aimed at students and teachers. On the other hand, they also point out that if a student uses AI against the guidelines, this can be considered intentional misuse. Although most guidelines do not mention how teachers should react when they notice a student has been misusing AI, the JYU Movi guideline for teachers instructs teachers to approach this kind of situation similarly to any other suspected plagiarism and to use their pedagogical judgment. In this process, teachers should consider the student's understanding of the issue and direct them to further information if necessary. Thus, the teacher is seen as an authority who can start actions but is obligated to reconsider the situation from the student's perspective. This example also shows how the guidelines present the understanding of the ethical use of AI tools as a cooperation between students and teachers.

Data Privacy, Intellectual Property, Technical Robustness and Safety. The two principles, privacy and data governance, and technical robustness and safety are discussed in this chapter because the themes partially overlap in the guidelines.

Firstly, all universities had a strong emphasis on data privacy. In Sweden, AI can be used within the university's framework. For example, students' material cannot be put into Generative AI without their permission. Teachers cannot force students to put personal information into AI tools that the university has not approved (Sweden, Finland). In Iceland, each institution explicitly requires that AI tools be used in a manner that complies with existing data protection laws, ensuring the security and confidentiality of personal and sensitive information. Their documents reflect a broad recognition of the risks associated with data breaches or improper data handling in AI applications and a proactive approach to mitigating these risks within academic settings. In Finland, there is a variation in how guidelines are presented, ranging from guidelines explaining in detail

what data protection and privacy rules mean to more generic and implicit presentations of guidelines. The Danish guidelines clearly state that Generative AI should not include sensitive, confidential, or personal data. The ban also applies to other people's personal data (e.g., names, gender, age, political affiliations, email addresses, or other personally identifiable information).

Additionally, only publicly available or legally disclosed information could be given to the Generative AI. Caution is also urged when personal data is processed, including information provided during user registration and less visible data such as location, browsing history, device, and browser details, and transferring personal data to countries outside the EU and General Data Protection Regulation (GDPR). For example, in Norwegian universities, an employee is not allowed to exchange data with American companies that do not have a data agreement with the university or are aligned with European privacy standards. At NTNU and UiO, many of their already implemented tools and software, like Zoom, Panopto, WeVideo, and LMS, haveoptional AI-powered features, which are not being used or made available for now. First, before making these new features and add-ons available, they must be investigated and approved via a thorough and extensive DPIA (Data Protection Impact Assessment) process. A data protection impact assessment (DPIA) aims to ensure that the system protects the privacy of data subjects – the people whose personal data will be collected, held, or processed.

In addition, Norwegian universities are still trying to adjust to the new EU AI-ACT to leverage their own/Governmental regulations and how this will inflict with or include DPIA, security levels, and storage regulations of various types of data.

Additionally, individuals are fully responsible for ensuring compliance with GDPR and must refrain from disclosing confidential information, particularly data derived from companies or other restricted sources (DK).

Intellectual property rights are mentioned in almost every guideline. Although not specified in all guidelines (e.g., in Sweden), they are implicitly mentioned when students' responsibilities and rights to their work are explained. In Iceland, for example, all existing rules on academic integrity, proper citation, and the acknowledgment of sources remain applicable and must be strictly adhered to. In this context, two Icelandic universities and one Norwegian university also emphasize the importance of proper AI training. Also, teachers and students in Denmark are encouraged to familiarize themselves with security considerations in generative AI, including the varying data classification policies in different services. Moreover, one of the Finnish guidelines suggests that a data protection impact assessment must be conducted before personal data is used, and data protection notes must be regularly updated.

Transparency and Accountability. Since transparency and accountability are mentioned in several guideline texts and many principles overlap, the main results are presented in the same chapter. Accountability was mentioned in many guidelines, especially in the view of minimizing risks, which can occur either due to intentional or unintentional misuse of ai-based tools. For example, swedish guidelines highlighted the students' role in adhering to and understanding the rules and guidelines. Additionally, if the student does not follow or understand the rules and the teacher has not given sufficient instructions, it could be interpreted as unintentional or intentional misuse. Additionally, Danish universities mentioned the importance of rule compliance in the context of collaboration.

One Finnish guideline mentioned that the teacher needs to ensure that the students have the necessary AI skills. A Swedish guideline emphasized that the instructions need to be up-to-date. Two Norwegian documents highlighted the need to handle AI-generated information critically and with caution, as well as the importance of understanding the limitations and biases of AI. Moreover, Danish guidelines brought up the challenges of identifying generative AI in various platforms, raising awareness of how different tools can or cannot be used in academic work.

The guidelines showed slightly more direct or indirect mentions of transparency than other ethical principles. Almost every document pointed out that the university requires ethical use of AI and integrity, at least at a general level. Some documents (e.g., in Finland) gave more detailed instructions on proceeding in practice. Most of the documents stated that the students and teachers need to be transparent in their work and, for example, say when AI is used and how it is used by offering clear guidelines. Furthermore, teachers' obligation to communicate and discuss the ethical use of AI with students was mentioned. There were also warnings, for instance, in the Finnish, Swedish, and Icelandic documents, about what would happen if AI was misused. The Icelandic guidelines mentioned the potential failure of an assignment or a course, formal warnings, and possible expulsion, emphasizing strict adherence to academic integrity and ethical use of AI. An example of misuse can be using AI in exams, which is forbidden, except with the teacher's consent, or the distribution of wrongful or defamatory information, as pointed out by the Icelandic guidelines. The deliberate AI misuse disciplinary actions were not mentioned in detail in the Norwegian guidelines, although the importance of avoiding cheating and plagiarism was expressed. Unintentional misuse was not explicitly mentioned in any document.

To summarize, the following examples have been mentioned in the guidelines on how to minimize the risks and misuse of AI (DK = Denmark, FI = Finland, IS = Iceland, NO = Norway, SWE = Sweden):

- Data protection impact assessment (FI)
- The teacher's obligatory role is to increase student knowledge and understanding (NO, SWE, FI)
- Proper citation practices to avoid plagiarism (NO, FI, IS)
- Institutional evaluation and verification processes (NO)
- Common agreement within a group on how to use AI before the project to avoid misunderstandings and conflicts (DK)
- Guidelines and clear instructions for students and teachers about misuse, exam questions, etc. (IS, NO, FI)
- The tools developed and offered by the university need to be as understandable as possible (FI)
- The technical operations must be communicated actively (FI)
- Formal warnings (IS)

6 Discussion

This study asks what recurring themes are present in the AI policies, frameworks, and guidelines established by Nordic universities. Through qualitative content analysis and a theory-driven content analysis of AI policy documents and guidelines across 12 universities in Denmark, Finland, Iceland, Norway, and Sweden were revealed regarding how AI is governed in Nordic higher education institutions: the use of AI in higher education and ethical principles of AI in higher education guidelines. These themes reflect the opportunities and challenges associated with AI adoption in Nordic higher education, highlighting shared values and governance approaches across the region.

6.1 Use of Artificial Intelligence in Higher Education

Integrating AI into higher education in Nordic universities is primarily focused on teaching, learning, research, and institutional management. Policies and guidelines across institutions recognize AI's transformative potential while emphasizing the need for responsible and ethical adoption.

Teaching and Learning. AI is increasingly utilized to support personalized learning, enhance student engagement, and improve teaching methodologies. Universities emphasize the necessity of AI literacy among both educators and students to ensure informed usage. Policies often require educators to specify AI use guidelines in course syllabi, granting them the authority to determine how AI can or cannot be used in their courses. Transparency in AI use is a key principle, with students being required to disclose AI-generated content where applicable. Some institutions also stress AI's role in fostering creativity, encouraging students to critically engage with AI tools as aids rather than replacements for independent thinking. Despite AI's potential, concerns regarding academic integrity remain central. Many policies outline the risks of plagiarism and unauthorized AI-assisted work, stressing that AI should not be used to substitute human reasoning and effort. Some universities, particularly in Finland and Sweden, frame AI literacy as a crucial work-life skill, encouraging students to develop competency in evaluating and responsibly incorporating AI into their academic work.

Research and Management. AI's role in research is acknowledged primarily in graduate studies, which are used for brainstorming, drafting research questions, and enhancing the writing process. However, universities caution against over-reliance on AI-generated content, reinforcing the expectation that students and faculty critically assess AI outputs. In Finland and Iceland, students are required to disclose AI usage in research projects, and in some cases, specific evaluation rubrics are provided to assess AI's role in academic work. Institutional management policies highlight AI's application in data analytics, decision-making, and administrative processes. While AI is recognized as a tool for improving efficiency, universities emphasize the importance of maintaining human oversight to ensure fairness and accountability in institutional governance.

Implementation and Training. The successful integration of AI in higher education is contingent on structured training and implementation strategies. Policies in Nordic universities stress the importance of equipping educators and students with the necessary knowledge and skills to navigate AI applications effectively. Dedicated AI literacy

programs have been introduced in Norway and Finland, and structured courses and training materials have been made available. denmark has also initiated pilot programs to train faculty in AI competencies, ensuring they can incorporate AI tools into their teaching methodologies. Overall, while Nordic universities acknowledge the benefits of AI in enhancing education and institutional management, their policies emphasize ethical responsibility, transparency, and human oversight. AI is positioned as a supportive tool rather than a substitute for human judgment, reinforcing the need for continuous learning and adaptation in response to evolving AI technologies.

6.2 Ethical Principles of AI in University Guidelines

Almost all ethical guidelines in this study were at a general level and aimed at minimizing risks and challenges. Based on several studies, this is the case also in other guidelines [27]. Moreover, the principles mentioned in the guidelines tended to be overlapping and ambiguous. They did not include any context or explain how to implement the rules in practice like in the previous studies [27, 28, 30]. However, there were examples of enhancing accountability by minimizing risks, such as offering clear guidelines for teachers and students.

Universities across the Nordic countries emphasize key ethical principles guiding AI use in higher education. These include human agency and oversight, data privacy and intellectual property, transparency and accountability, and responsible AI implementation.

Human Agency and Oversight. Institutions stress that AI should serve as a tool to support, rather than replace, human decision-making in education. Universities encourage students and faculty to critically assess AI-generated content, aligning with broader goals of fostering independent thinking and ethical use. Nevertheless, the results imply that in policies and guidelines, students and teachers are seen as actors who work together to understand how AI tools can be used safely, according to laws, and ethically.

Data Privacy, Intellectual Property, Technical Robustness and Safety. Compliance with data protection regulations, including GDPR, is a key concern. Policies mandate that AI tools be used responsibly, protecting sensitive data and maintaining intellectual property rights.

Transparency and Accountability. Many guidelines require students and faculty to disclose the use of AI in academic work. Clear attribution practices are stressed to prevent academic misconduct like plagiarism or the misrepresentation of AI-assisted work.

7 Conclusion and Recommendations

The recurring themes in AI policies across Nordic universities reflect a shared commitment to leveraging AI responsibly while maintaining ethical, pedagogical, and institutional integrity. By prioritizing ethical considerations, transparency, and education, Nordic universities are shaping AI policies that support innovation and safeguard the core principles of higher education in an AI-driven era.

Based on the thematic analysis, the ethical principles of the guidelines mainly focused on data privacy and transparency, both of which also included misuse and disciplinary actions. Only the theory-driven content analysis revealed that almost all guidelines had a strong human agency perspective, which is even more important to consider in the future as automation and artificial intelligence use increases. A rather serious and worrying finding is that the guidelines were missing considerations on the principles of social and environmental well-being, diversity, non-discrimination, and fairness. This omission is concerning, as these principles are increasingly important today, given the growing significance of human rights and environmental issues in our society and within academic communities. Since students, teachers, and researchers are all struggling to keep up with the speed of AI development, they are also learning together how AI can be used safely and ethically. When students and teachers discuss ethical AI use principles, they work together to make shared meanings of AI use in HEIs and academic contexts. While they negotiate together on how and why AI tools can be used in their studies, they also negotiate the boundaries of their relationship, which is deeply connected to their societal and academic well-being. Another crucial result, like in the previous studies, was that the guidelines did not explain how the principles are followed in practice.

Acknowledgements. This study was conducted in a cross-Nordic collaboration of the EDUCERE NORDIC network funded by The Swedish Research Council (Vetenskapsrådet) 2023–2025, Research Network Grant 202-04469.

Disclosure of Interests. The authors declare that they have no competing interests relevant to this article's content.

Declaration of Generative AI and ICT Tools. This study employed various Artificial Intelligence (AI) and Information and Communication Technology (ICT) tools to facilitate research and analysis. scispace, semantic scholar, and scopus were utilized for reference searches. copilot enterprise assisted in the analysis of Norwegian documents, while Atlas.ti was used for examining icelandic documents. additionally, experiments were conducted on the Icelandic documents using ChatGPT for analysis, followed by a comparative evaluation, which revealed ChatGPT's limitations in reliability. Grammarly was used for text editing and proofreading. after using those tools, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

Appendix

The documents analyzed are listed in an Excel file titled 'Countries_Universities_Documents_Analysed.xlsx,' which is available online. It can be accessed at the following URL: https://github.com/sibbajoa/EDUCERE.

References

1. Raj, A.: Artificial intelligence. Int. J. Res. Appl. Sci. Eng. IJRASET, **12**(XI) (2024). https://doi.org/10.22214/ijraset.2024.64695

2. Shrivastava, A., Pandey, A., Singh, N., Srivastava, S., Srivastava, M., Srivastava, A.: Artificial Intelligence (AI): evolution, methodologies, and applications. Int. J. Res. Appl. Sci. Eng. Technol. IJRASET **12**(IV) (2024). https://doi.org/10.22214/ijraset.2024.61241
3. Marengo, A., Pange, P.J.: Envisioning general AI in higher education: transforming learning paradigms and pedagogies. In: Guralnick, D., Auer, M.E., Poce, A. (eds.) Creative Approaches to Technology-Enhanced Learning for the Workplace and Higher Education. TLIC 2024. Lecture Notes in Networks and Systems, vol. 1150. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-72430-5_28
4. Del Mundo, M.A., et al.: Discourse analysis on experience-based position of science, mathematics, and Tech-Voc educators on generative AI and academic integrity. Environ. Soc. Psychol. **9**(8) (2024). https://doi.org/10.59429/esp.v9i8.3028
5. UNESCO, F., Miao, W., Holmes, H., Ronghuai,, Hui, Z.: AI and education: guidance for policy-makers (2021). https://doi.org/10.54675/PCSP7350
6. Luckin, R., Holmes, W., Griffiths, M., Forcier, L.B.: Intelligence Unleashed: An argument for AI in Education. https://www.pearson.com/corporate/about-pearson/what-we-do/innovation/smarter-digital-tools/intelligence-unleashed.html. Accessed 15 February 2025
7. Soule, K.: What is generative AI? IBM Research. https://research.ibm.com/blog/what-is-generative-AI. Accessed 14 February 2025
8. Selwyn, N.: On the Limits of Artificial Intelligence (AI) in Education. Nord. Tidsskr. Pedagog. Og Krit. **10**(1), Art. no. 1 (2024). https://doi.org/10.23865/ntpk.v10.6062
9. Tzirides, A.O., et al.: Generative AI: implications and applications for education (2023). arXiv:2305.07605. https://doi.org/10.48550/arXiv.2305.07605
10. Cope, B., Kalantzis, M.: A multimodal grammar of artificial intelligence: measuring the gains and losses in generative AI. Multimod. Soc. **4**(2), 123–152 (2024). https://doi.org/10.1177/26349795231221699
11. Ayeni, O.O., et al.: AI in education: A review of personalized learning and educational technology. GSC Adv. Res. Rev. **18**(2), Art. no. 2 (2024). https://doi.org/10.30574/gscarr.2024.18.2.0062
12. Global Initiatives Towards Regulatory Frameworks for Artificial Intelligence (AI) in Higher Education|Digital Government: Research and Practice. Accessed 14 February 2025. https://doi.org/10.1145/3672462
13. Al Daraai, S.B., Al Maqrashi, M., Al Zakwani, M., Al Shaikh, Z.: Integrating AI in higher education: applications, strategies, ethical considerations. In: Utilizing AI for Assessment, Graduations, and Feedback in Higher Education, pp. 189–211. IGI Global (2024). https://doi.org/10.4018/979-8-3693-2145-4.ch008
14. Al Harrasi, N., El Din, M.S.: Utilizing AI for assessment, grading, and feedback in higher education. In: Utilizing AI for Assessment, Grading, and Feedback in Higher Education. IGI Global (2024). https://doi.org/10.4018/979-8-3693-2145-4
15. Sucena, A., Cabrilo, S., Durst, S.: The integration of ChatGPT in higher education: a study of future prospects. In: Proceedings of 25th European Conference on Knowledge Management, vol. 25, no. 1 (2024). https://doi.org/10.34190/eckm.25.1.2341
16. Krpan, D., Mladenovic, S.: Unveiling the black box: Understanding AI in education. European Digital Education Hub (2025). https://eceuropaeu.sharepoint.com/teams/GRP-EACext-DigitalEducationHub/Shared%20Documents/Forms/Public.aspx?id=%2Fteams%2FGRP%2DEACext%2DDigitalEducationHub%2FShared%20Documents%2FReading%20Corner%2FShort%20Informative%20Papers%2FUnderstanding%20AI%20in%20Education%2Epdf&parent=%2Fteams%2FGRP%2DEACext%2DDigitalEducationHub%2FShared%20Documents%2FReading%20Corner%2FShort%20Informative%20Papers
17. Zeer, M., Siaj, R.W., Ghannam, J.A., Kanan, M.: Ethics of artificial intelligence in university education. In: Presented at the 2023 2nd International Engineering Conference on Electrical,

Energy, and Artificial Intelligence (EICEEAI 2023), pp. 635–638. Institute of Electrical and Electronics Engineers Inc. (2023). https://doi.org/10.1109/EICEEAI60672.2023.10590285
18. Alqahtani, N., Wafula, Z.: Artificial intelligence integration: pedagogical strategies and policies at leading universities. Innov. High. Educ. (2024). https://doi.org/10.1007/s10755-024-09749-x
19. Felzmann, H., Villaronga, E.F., Lutz, C., Tamò-Larrieux, A.: Transparency you can trust: transparency requirements for artificial intelligence between legal norms and contextual concerns. Big Data Soc. **6**(1), 2053951719860542 (2019). https://doi.org/10.1177/2053951719860542
20. Floridi, L., Cowls, J.: A unified framework of five principles for AI in society. Harv. Data Sci. Rev. **1**(1) (2019). https://doi.org/10.1162/99608f92.8cd550d1
21. Raina, A., Mundra, K., Prasad, P., Mishra, S.C.: Fostering ethics in AI: perceptions from the Indian AI curriculum. In: Presented at the 31st International Conference on Computers in Education, ICCE 2023 – Proceedings, Shih, J.L., et al. (eds.) Asia-Pacific Society for Computers in Education, pp. 609–616 (2023). https://www.scopus.com/inward/record.uri?eid=2-s2.0-85181770489&partnerID=40&md5=699a575e490a9676b26985dbac1ae652
22. An, Q., Yang, J., Xu, X., Zhang, Y., Zhang, H.: Decoding AI ethics from users' lens in education: a systematic review. Heliyon **10**(20) (2024). https://doi.org/10.1016/j.heliyon.2024.e39357
23. Jobin, A., Ienca, M., Vayena, E.: The global landscape of AI ethics guidelines. Nat. Mach. Intell. **1**(9), 389–399 (2019). https://doi.org/10.1038/s42256-019-0088-2
24. Bryman, A.: Social Research Methods. Oxford University Press (2016)
25. Recommendation of the Council on Artificial Intelligence. https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449. Accessed 14 February 2025
26. Hagendorff, T.: The Ethics of AI Ethics: An Evaluation of Guidelines|Minds and Machines. https://doi.org/10.1007/s11023-020-09517-8. Accessed 14 February 2025
27. Morley, J., Floridi, L., Kinsey, L., Elhalal, A.: From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. Sci. Eng. Ethics **26**(4), 2141–2168 (2020). https://doi.org/10.1007/s11948-019-00165-5
28. Cath, C.: Governing artificial intelligence: ethical, legal and technical opportunities and challenges. Philos. Trans. R. Soc. Math. Phys. Eng. Sci. (2018). https://doi.org/10.1098/rsta.2018.0080
29. Waelen, R.: Why AI ethics is a critical theory. Philos. Technol. **35**(1), 9 (2022). https://doi.org/10.1007/s13347-022-00507-5
30. Prem, E.: Approaches to ethical AI. In: Werthner, H., et al. (eds.) Introduction to Digital Humanism. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-45304-5_15
31. Funa, A.A., Gabay, R.A.E.: Policy guidelines and recommendations on AI use in teaching and learning: a meta-synthesis study. Soc. Sci. Human. Open **11** (2025). https://doi.org/10.1016/j.ssaho.2024.101221
32. Fabijanić Gagro, S.: Artificial intelligence in education – current challenges. Anali Pravnog Fak. U Beogr. **72**(4), 725–747 (2024). https://doi.org/10.51204/Anali_PFBU_24405A
33. Sethy, A., Shaik, N., Yadavalli, P.K., Anandaraj, S.P.: AI: issues, concerns, and ethical considerations. In: Toward Artif. Gen. Intell.: Deep Learn., Neural Networks, Gener. AI, De Gruyter, pp. 189–211 (2023). https://doi.org/10.1515/9783111323749-009
34. Zhan, Z., Rahman, A.B.M.M., Paul, A., Weiguo, F., Entian, J.: Evaluating the adaptability of AI technology in higher education. In: Li, E.Y., Chau, P.Y.K., Cheung, C.M.K. (eds.) Presented at the Proceedings of the International Conference on Electronic Business (ICEB), International Consortium for Electronic Business, pp. 639–644 (2024). https://www.scopus.com/inward/record.uri?eid=2-s2.0-85213382745&partnerID=40&md5=d596d77d3d569feb68f3ae0e4e87d60f

35. McGrath, C., Åkerfeldt, A.: Educational technology (EdTech): unbounded opportunities or just another brick in the wall? In: Digital Transformation and Public Services. Routledge (2019)
36. Beshr, B., Ateeq, A., Ateeq, R.A., Alaghbari, M.A.. AI in academia: pros and cons of integrating artificial intelligence in universities. In: AlDhaen, E., Braganza, A., Hamdan, A., Chen, W. (eds.) Business Sustainability with Artificial Intelligence (AI): Challenges and Opportunities. Studies in Systems, Decision and Control, vol. 566. Springer, Cham (2025). https://doi.org/10.1007/978-3-031-71318-7_31
37. Gordienko, O., Bagrationi, K.: Global integration of artificial intelligence in higher education sector: a bibliometric analysis. In: Obermayer, N., Bencsik, A. (eds.) Presented at the Proceedings of the European Conference on Knowledge Management, ECKM, pp. 270–276. Academic Conferences and Publishing International Limited (2024). https://doi.org/10.34190/eckm.25.1.2787
38. Azevedo, L., Mallinson, D.J., Wang, J., Robles, P., Best, E.: AI policies, equity, and morality and the implications for faculty in higher education. Pub. Integr. (2024). https://doi.org/10.1080/10999922.2024.2414957
39. Kim, D., Wu, J.: Artificial intelligence in higher education: examining the AI policy landscape at U.S. institutions. In: Presented at the Proceedings of 25th Annual Conference on Information Technology Education, SIGITE 2024, pp. 68–73. Association for Computing Machinery, Inc. (2024). https://doi.org/10.1145/3686852.3687076
40. NOKUT. Studiebarometeret 2024 – Hovedtendenser (2025). https://www.nokut.no/globalassets/studiebarometeret/2025/studiebarometeret-2024-hovedtendenser_1-2025.pdf
41. Den fællesoffentlige digitaliseringsstrategi. https://www.digmin.dk/digitalisering/mere-om-digitalisering/den-faellesoffentlige-digitaliseringsstrategi. Accessed 14 February 2025
42. Ny strategisk indsats skal bane vej for kunstig intelligens i Danmark. https://www.digmin.dk/digitalisering/nyheder/nyhedsarkiv/2024/dec/ny-strategisk-indsats-skal-bane-vej-for-kunstig-intelligens-i-danmark. Accessed 14 February 2025
43. Finland 2024 Digital Decade Country Report|Shaping Europe's Digital Future. https://digital-strategy.ec.europa.eu/en/factpages/finland-2024-digital-decade-country-report. Accessed 14 February 2025
44. Regulation – EU – 2024/1689 – EN – EUR-Lex. https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng. Accessed 14 February 2025
45. AI Regulation – Työ- ja elinkeinoministeriö – YJA, Työ- ja elinkeinoministeriö. https://tem.fi/ai-regulation. Accessed 15 Februry 2025
46. Measuring Digital Development – Facts and Figures 2024. ITU. https://www.itu.int/hub/publication/d-ind-ict_mdd-2024-4/. Accessed 14 February 2025
47. Digital Iceland. https://island.is/en/o/digital-iceland. Accessed 14 February 2025
48. Stefna Íslands um gervigreind.pdf. https://www.stjornarradid.is/library/01--Frettatengt---myndir-og-skrar/FOR/Fylgiskjol-i-frett/Stefna%20%c3%8dslands%20um%20gervigreind. Accessed 14 February 2025
49. Head start for Icelandic. https://www.government.is/diplomatic-missions/embassy-article/2023/03/14/Head-start-for-Icelandic/. Accessed 14 February 2025
50. OpenAI. "Government of Iceland," OpenAI
51. Aðgerðaáætlun um gervigreind nóvember 2024.pdf. https://www.stjornarradid.is/library/04-Raduneytin/Haskola---idnadar--og-nyskopunarraduneytid/A%C3%B0ger%C3%B0a%C3%A1%C3%A6tlun%20um%20gervigreind%20n%C3%B3vember%202024.pdf. Accessed 14 February 2025
52. AI Act|Shaping Europe's Digital Future. https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai. Accessed 11 February 2025

53. K. Moderniseringsdepartementet. EUs forslag til regulering av KI – Norges posisjon. https://www.regjeringen.no/no/tema/statlig-forvaltning/it-politikk/KI-strategi/nasjonal-posisjon-pa-ki-regulering/id2870299/. Accessed 11 February 2025
54. J.beredskapsdepartementet. Norge undertegnet Europarådskonven-sjonen om kunstig intelligens. https://www.regjeringen.no/no/aktuelt/norge-undertegnet-europaradskonvensjonen-om-kunstig-intelligens/id3052724/. Accessed 11 February 2025
55. Nordic innovation Ethical AI Expert Group. Nordic Innovation. https://www.nordicinnovation.org/2024/nordic-ethical-ai-expert-group. Accessed 24 January 2025
56. OECD Digital Government Index|Society Indicators. OECD Going Digital Toolkit. https://goingdigital.oecd.org/en/indicator/58. Accessed 10 February 2025
57. Shaping Norway's Digital Future, OECD. https://www.oecd.org/en/publications/shaping-norway-s-digital-future_d3af799c-en.html. Accessed 10 February 2025
58. D. forvaltningsdepartementet. The Digital Norway of the Future. https://www.regjeringen.no/en/dokumenter/the-digital-norway-of-the-future/id3054645/. Accessed 31 January 2025
59. M. of L. G. and Modernisation. The National Strategy for Artificial Intelligence. https://www.regjeringen.no/en/dokumenter/nasjonal-strategi-for-kunstig-intelligens/id2685594/. Accessed 31 January 2025
60. Home – NORA – Norwegian Artificial Intelligence Research Consortium. https://www.nora.ai/index.html. Accessed 10 February 2025
61. Norwegian Research Center for AI Innovation (NorwAI) – NTNU. https://www.ntnu.edu/norwai/. Accessed 10 February 2025
62. Norwegian Open AI Lab – NTNU. https://www.ntnu.edu/ailab. Accessed 10 February 2025
63. M. of Finance. Norwegian Agency for Public and Financial Management (DFØ). https://www.regjeringen.no/en/dep/fin/about-the-ministry/subordinateagencies/the-norwegian-government-agency-for-fina/id270409/. Accessed 10 February 2025
64. Skyreisen – Markedsplassen for skytjenester. https://tjenester.markedsplassen.anskaffelser.no/veiledning/skyreisen. Accessed 11 February 2025
65. Tid for generativ KI i sandkassa. Datatilsynet. https://www.datatilsynet.no/aktuelt/aktuelle-nyheter-2023/tid-for-sprakmodeller-i-sandkassa/. Accessed 11 February 2025
66. Copilot prosjekt – NTNU. https://www.ntnu.no/adm/it/copilot. Accessed 24 January 2025
67. The Swedes and the Internet. Svenskarna och internet. https://svenskarnaochinternet.se/english/. Accessed 14 February 2025
68. The Swedish Higher Education Act (1992:1434). UHR.se. https://www.uhr.se/en/start/laws-and-regulations/Laws-and-regulations/The-Swedish-Higher-Education-Act/. Accessed 14 February 2025
69. The Higher Education Ordinance (1993:100). UHR.se. https://www.uhr.se/en/start/laws-and-regulations/Laws-and-regulations/The-Higher-Education-Ordinance/. Accessed 14 February 2025
70. Ethics guidelines for trustworthy AI|Shaping Europe's digital future. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai. Accessed 14 February 2025

A Property Checklist for Evaluating the Student Experience with Consideration of Cultural Aspects

Nicolás Matus[1,2(✉)], Cristian Rusu[1], and Federico Botella[2]

[1] Pontificia Universidad Católica de Valparaíso, Av. Brasil 2241, 2340000 Valparaíso, Chile
nicolas.matus.p@mail.pucv.cl, cristian.rusu@pucv.cl
[2] Universidad Miguel Hernández de Elche, Avenida de la Universitat s/n, 03202 Elche, Spain
federico@umh.es

Abstract. The demand for higher-quality education has become a reality in today's global landscape, driven by rapid societal change and technological innovation. UNESCO's Sustainable Development Goals place improving education at the forefront, emphasizing the need to raise standards, reduce inequalities, and foster equity in access and delivery. Achieving these goals requires a comprehensive, cross-sectoral strategy that applies specialized tools to address complex challenges within educational contexts.

A central element in meeting these objectives is enhancing the Student eXperience (SX), an area increasingly recognized in higher education. The SX encompasses all physical, cognitive, and emotional interactions that students—both current and prospective—have with Higher Education Institutions (HEIs). Similar to the Customer eXperience (CX) area, evaluating the SX entails considering not only academic and infrastructural provisions but also psycho-emotional support, social integration, and availability of resources. Addressing the SX must account for diverse cultural backgrounds, as these can substantially influence student interactions and expectations. Incorporating models like Hofstede's cultural dimensions into assessment methods allows HEIs to gain deeper insight into how varying cultural factors shape student perceptions and behaviors.

In this context, we propose a property checklist as part of a structured evaluation methodology for SX. Designed as an initial diagnostic instrument, the checklist guides evaluators through a systematic review of key properties, criteria, and components influencing the student journey. By illuminating specific areas for improvement, this tool opens opportunities for targeted interventions that cater to the unique needs of diverse student populations. Ultimately, such a multifaceted approach fosters higher-quality education, mitigates inequalities, and advances equity and diversity across educational services.

Keywords: Student Experience · Customer eXperience · Culture · Property Checklist · Evaluation · Methodology · Diversity

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 219–235, 2025.
https://doi.org/10.1007/978-3-031-93539-8_15

1 Introduction

Considering the pivotal role that perceived quality plays in Higher Education Institutions (HEIs) for accreditation and market positioning, it becomes clear that investing in mechanisms capable of addressing students' expectations and perceptions [1] and evaluating and managing their educational experiences is paramount. Although a broad range of tools is currently employed to assess the quality of higher education services, these instruments are often generic and fail to account for the holistic student experience [2]. Furthermore, they typically overlook crucial elements of students' backgrounds, such as cultural factors.

In response to these limitations, we propose the development of a property checklist as part of a more comprehensive, holistic methodology for evaluating the Student eXperience (SX). The primary purpose of this checklist is to assist evaluators—those responsible for implementing the methodology—in planning and prioritizing tasks and activities in a logical sequence while selecting the most appropriate methods to meet the specific needs of HEIs.

This paper presents the theoretical framework that underpins our proposal, outlines the process through which a preliminary version of the checklist was developed, and discusses the key considerations and conclusions derived from this process as a whole. Future research will focus on evaluating, validating, and consolidating a definitive checklist for full integration into our evaluation methodology.

2 Theoretical Background

2.1 User eXperience

User eXperience (UX) shapes how individuals interact with a product, system, or service. As ISO standard 9241–210 defines, UX is the "person's perceptions and responses resulting from the use and/or anticipated use of a product, system, or service" [3]. This concept is particularly relevant to our study, as students engage with various services, products, and systems within HEIs, effectively positioning them as users or clients of these offerings.

UX encompasses a broad spectrum of elements that influence students in higher education, including emotions, cognitive processes, personal preferences, perceptions, and physiological and psychological responses. Given that students interact with multiple educational systems and services throughout their academic journey, a comprehensive analysis of their experiences is essential. From a Customer eXperience (CX) perspective, this entails examining their interactions across all touchpoints within the educational environment, ensuring a holistic understanding of how they engage with institutional products, services, and systems [4]. Furthermore, an in-depth assessment of UX/CX in HEIs can reveal critical insights into students' expectations, satisfaction levels, and potential challenges they encounter. By identifying and addressing these factors, institutions can enhance the overall learning experience, optimize service delivery, and foster a more supportive academic environment.

2.2 Customer eXperience

The concept of CX, as a theoretical extension of UX, directly connects the customer—in the context of this proposal, the student—with the products and services they utilize, emphasizing, in particular, their reactions and perceptions. Hill et al. (2007) define CX as "the physical and emotional experiences that occur through interactions with the products and/or services offered by a brand/company from the point of first direct and conscious contact, considering the total journey, until the post-consumption stage" [5]. In addition, several authors have suggested that CX comprises six key dimensions [6]: (i) emotional, (ii) sensory, (iii) cognitive, (iv) pragmatic, (v) lifestyle, and (vi) relational.

From an academic standpoint, understanding CX is crucial in traditional customer-facing scenarios and educational contexts where students engage with various digital systems, services, and platforms. Our proposal, therefore, adopts a Human-Computer Interaction (HCI) approach to CX, drawing upon established evaluation and analysis methods from this domain. In doing so, we examine client–system or IT product interactions, which are increasingly prevalent across various industries. These interactions are scrutinized through user-centered methods, usability assessments, and other HCI-based techniques to gain deeper insight into the user's overall experience.

Furthermore, effective CX management has been widely recognized for its potential to confer competitive brand advantages by improving customer satisfaction [7]. This advantage applies equally in educational and corporate settings, where successfully aligning product or service offerings with user expectations can significantly enhance perceived value, loyalty, and overall satisfaction. Consequently, integrating HCI principles into CX management ensures a more holistic and user-focused approach, thereby fostering meaningful and impactful experiences.

2.3 Student eXperience

SX has been defined in various ways and lacks a universally accepted definition. It is crucial to highlight that numerous scholars have emphasized the ubiquity of the student experience, which frequently extends beyond purely academic environments [8, 9]. Nevertheless, the concept can be analyzed as a specific instance of CX, wherein the customers are the students who utilize the products and services provided by HEIs. From this standpoint, SX has been defined as "all the physical and emotional perceptions and reactions that a student or prospective student experiences in response to interaction with the products, systems, or services provided by an HEI, and interactions with people related to the academic field, both inside and outside of the academic space" [10].

The significance of HEIs investing in SX is underscored by the alignment with the UNESCO Sustainable Development Goals (SDGs) addressed by our methodology [11, 12]. SDGs such as Quality Education, Gender Equality, and Reduced Inequalities are directly targeted to enhance student satisfaction and the perceived quality of HEIs.

Furthermore, our approach strongly emphasizes the cultural aspects of the students being evaluated. A comprehensive understanding of SX allows institutions to tailor their policies and initiatives more effectively, fostering inclusive, student-centered environments that promote academic success and personal development. Additionally, considering the increasing globalization of higher education, addressing SX holistically can contribute to enhancing institutional reputation and competitiveness on a broader scale.

2.4 Culture

Culture is a fundamental element of societies that, when analyzed, allows for detecting differences in behavior, ethics, and values. Hofstede (2010) defined culture as "the collective programming of the mind that distinguishes the members of one group or category of people from another" [13]. In the specific context of SX, analyzing students' cultural backgrounds can help us assess the degree of satisfaction or dissatisfaction they experience throughout their university life based on their nationality. This analysis is particularly significant for refugee students, children of migrants, and exchange students.

We have adopted the national culture dimensions proposed by Hofstede to develop our evaluation methodology. This model encompasses six dimensions: (i) Power Distance (PD), (ii) Uncertainty Avoidance (UA), (iii) Individualism/Collectivism (IDV), (iv) Masculinity/Femininity (MAS), (v) Long/Short Term Orientation (LTO), and (vi) Indulgence/Restraint (IVR). The rationale for selecting Hofstede's cultural model [13, 14] over other models is its extensive application and validation by numerous researchers. Furthermore, many cultural models developed after Hofstede's share common elements [15].

It is worth noting that the Masculinity vs. Femininity dimension has recently been renamed to Motivation towards Achievement and Success [16]. This change was prompted by concerns raised by some researchers who argued that associating specific values and behaviors inherent to a culture with socially constructed gender roles reflected a biased perspective, one that perpetuates gender stereotypes [17]. The revision aims to provide a more neutral framework that better captures the cultural differences related to achievement and success without reinforcing outdated notions of gender. By reframing this dimension, scholars hope to promote a more inclusive understanding of cultural values that avoids conflating gender-specific traits with national cultural characteristics. This shift acknowledges the evolving discourse on gender and culture in social sciences, encouraging a more nuanced approach to intercultural studies.

3 Student eXperience Evaluation

To enhance the perceived quality of HEIs and improve student satisfaction and overall well-being—regardless of their cultural background—we have proposed a SX evaluation methodology that explicitly incorporates cultural factors inherent to students [18] (Fig. 1). This methodology is grounded in the SX-Cultural Model, which we developed by integrating our three-dimensional SX model with Hofstede's six-dimensional National Culture Model [19, 20].

As illustrated in Fig. 1, the methodology consists of five distinct stages. The first two stages form part of the diagnostic and preparation process, while the final two stages are dedicated to consolidation, analysis, and presentation of results. The central stage (III), which serves as the core execution phase, involves the application of evaluation methods aligned with the predefined objectives of the evaluation process.

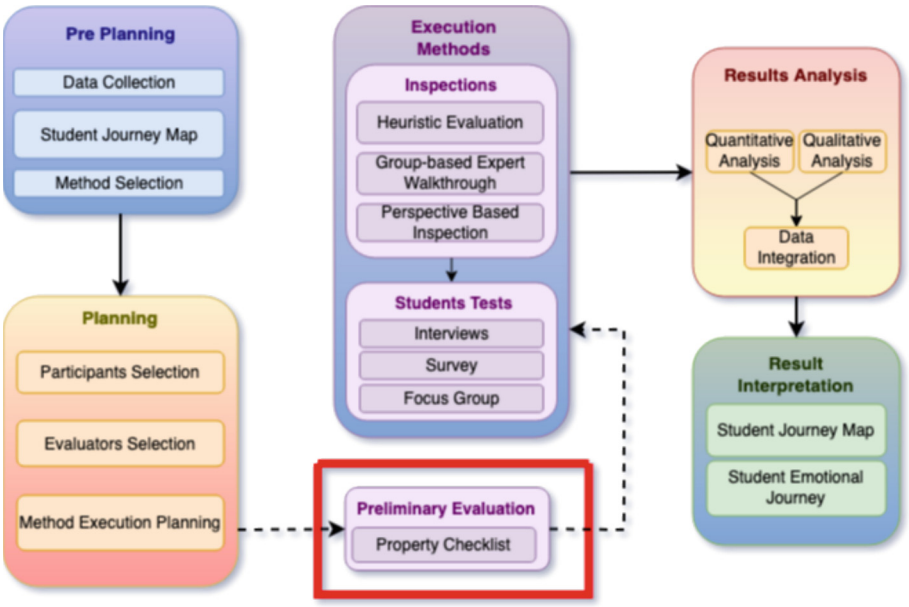


Fig. 1. Student Experience Evaluation Methodology Incorporating Cultural Factors.

Within this methodology stage, we propose the checklist, designed to serve as a preliminary evaluative method. This checklist must be administered before any other evaluation methods, as its primary function is to confirm that we have the requisite resources, participants, and infrastructure—initially identified in the pre-planning and planning stages—readily available for the effective implementation of subsequent inspection and evaluation methods.

Furthermore, we have extended the scope of the checklist to encompass key aspects of the data integration phase within the Results Analysis Stage, ensuring that all information gathered thus far is consolidated and adequately prepared for detailed analysis. It is essential to note that, depending on the specific objectives of the evaluation process, one or more of the proposed evaluation methods may be deemed unnecessary and can, therefore, be excluded. For instance, if our evaluation does not analyze students' interactions with particular products, services, or software systems, we may omit specific inspection methods and proceed directly to the student tests.

It is essential to highlight that, considering the theoretical foundation based on CX as a conceptual extension of UX, the holistic evaluation process of the student experience is approached with a strong HCI component. This emphasis is particularly evident in the execution stage, which includes usability inspections that may explicitly assess interactions with electronic products, systems, and software services. By integrating HCI principles, the evaluation framework ensures a comprehensive analysis of the technological and experiential dimensions that shape the student experience.

4 Related Works

Almeida et al. (2020) introduce a dedicated checklist for evaluating UX in IoT environments, demonstrating the viability of adapting comparable frameworks to the SX context [21]. Conceptually, SX can be regarded as an extension of CX and, by extension, UX; thus, broadening its scope to include SX should not pose significant methodological challenges. Moreover, the successful application of this IoT-focused evaluation tool reinforces the practicality of integrating checklists within a more comprehensive evaluation methodology, allowing for a systematic analysis of student interactions with software products, systems, and services in varied educational and technological settings within HEIs. In addition, Richardson et al. (2021) propose a Mobile Application User Experience Checklist, further illustrating the development of checklists in the broader field of software services interaction [22]. Notably, mobile technology functions as a pivotal communication enabler, which—when applied to SX—spans all three dimensions (Educational, Social, and Personal) [20]. Since students are regularly exposed to multiple communication channels, this environment introduces various cultural factors and dynamics that necessitate careful consideration to maintain an inclusive and supportive educational experience.

Beyond their widespread use in UX and HCI, checklists have proven to be effective in structuring evaluations by providing systematic criteria that enhance consistency and reliability. Their adaptability has been evident in diverse domains, ranging from usability assessments in software engineering to HCI. The transition from UX to SX evaluation is not only methodologically feasible but also advantageous, as SX inherently involves complex user interactions within educational ecosystems. The application of checklists allows for a structured, repeatable, and scalable approach to assessing student interactions with technology, ensuring that key experiential factors—such as accessibility, engagement, and overall satisfaction—are systematically analyzed.

Moreover, the novelty of applying checklists in an SX evaluation lies in their ability to capture multi-dimensional aspects of the student experience, which go beyond usability and interface considerations to encompass social, cultural, and emotional engagement with educational technologies. The structured nature of checklists facilitates their integration into both qualitative and quantitative research methodologies, enabling researchers and practitioners to gather comprehensive insights while maintaining methodological rigor. This structured approach is particularly beneficial in HEIs, where diverse student backgrounds and technological ecosystems demand a flexible yet consistent evaluation framework.

5 Preliminary Checklist Proposal

5.1 Checklist Development

Stufflebeam's Guidelines for Developing Checklists guided the development of this checklist—also known as the *Checklist Development Checklist (CDC)* (2000) [23]—which provides a comprehensive framework for the creation and refinement of evaluative instruments. The general model consists of twelve stages: (I) Focus the checklist task, (II) Make a candidate list of checkpoints, (III) Classify and sort the checkpoints, (IV)

Define and flesh out the categories, (V) Determine the order of categories, (VI) Obtain initial reviews of the checklist, (VII) Revise the checklist content, (VIII) Delineate and format the checklist to serve the intended uses, (IX) Evaluate the checklist, (X) Finalize the checklist, (XI) Apply and disseminate the checklist, and (XII) Periodically review and revise the checklist. This framework establishes a systematic and iterative process to ensure that the final instrument is comprehensive and rigorously aligned with its intended objectives.

The initial stages involved clearly defining the domain of interest and validating key merit criteria by engaging with relevant literature and consulting subject-matter experts. Once a broad set of descriptors was identified, they were systematically classified into logical categories. Each category was subsequently refined and contextualized with precise definitions, rationales, and considerations regarding potential overuse. Depending on user needs, the order of these categories may also be deliberately structured.

Following this, the checklist undergoes multiple rounds of review—first by potential users and experts—to identify and address any gaps or ambiguities. Revisions are then made based on the feedback received, after which further user engagement is conducted to finalize scoring and weighting systems tailored to the specific evaluation context. The checklist is finalized and disseminated once its applicability, clarity, and usability are confirmed. Mechanisms for user feedback and periodic updates are integrated to ensure its continued relevance and effectiveness over time.

In alignment with these principles and building upon the conceptual foundations of our SX-Cultural model [19, 20], a series of key activities were identified for integration into the checklist. These activities are organized into four conceptual phases (Table 1), each corresponding to one of the five categories within the broader evaluation methodology in which the checklist will be implemented (Fig. 1) [18]. Furthermore, the conceptualization of these phases was guided by the necessity of capturing context-specific and multidimensional data, ensuring that each thematic component is systematically addressed and analyzed in sufficient depth.

This process was carried out following the first five stages of the CDC, consolidating our preliminary proposal. It represents the initial phase of a three-stage process in the development and application of our checklist, which will subsequently be followed by an Evaluation and Validation phase (II), ultimately culminating in an iterative cycle of application and continuous improvement (III) (Fig. 2).

Table 1 provides a detailed overview of the checklist key activities, outlining explicit objectives tailored to each thematic area. This structured approach is designed to facilitate a rigorous and methodical evaluation process, thereby enhancing both the findings' reliability and the resulting recommendations' practical applicability.

By applying the checklist in Table 1, evaluators can systematically assess the SX across its three dimensions while incorporating cultural considerations derived from Hofstede's cultural model. This structured approach ensures that the resulting insights and recommendations are comprehensive and actionable.

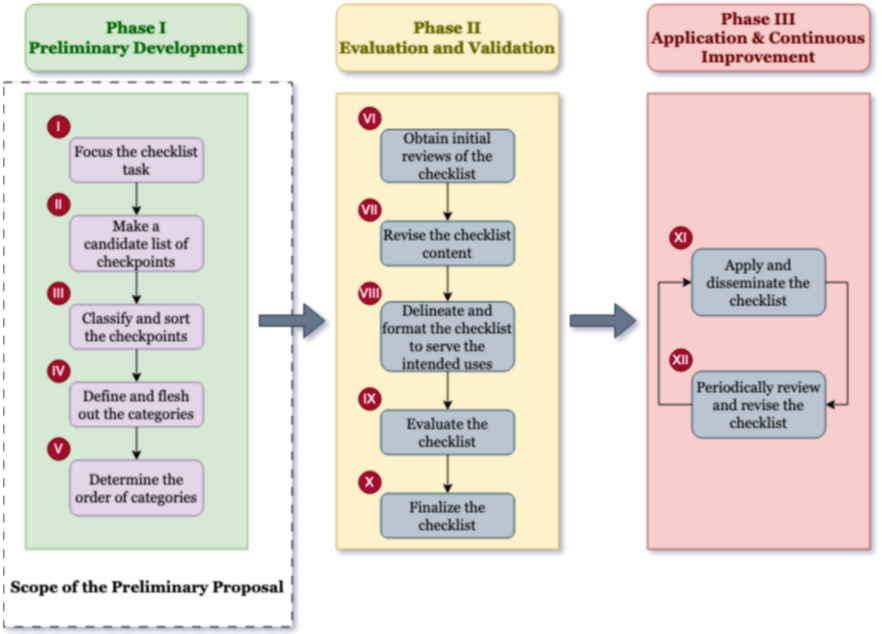


Fig. 2. Process of Development, Evaluation, Implementation, and Continuous Improvement of the Checklist.

Table 1. Proposed Checklist Items with Context and Goals.

Phase/Step	Key Activities	Goals
1. Preliminary Planning (Define Target Student Profiles)	- Identify the academic levels, cultural backgrounds, and learning contexts of the student groups involved - Consider Hofstede's six cultural dimensions	- Tailor the evaluation approach to diverse student populations - Ensure cultural sensitivity and inclusiveness - Align with broader institutional or program-specific goals
1. Preliminary Planning (Establish Evaluation Scope & Objectives)	- Determine whether the evaluation focuses on web-based platforms, software, or blended-learning environments - Outline specific timelines for data collection and analysis (e.g., deadlines for inspections, student testing, reporting) - Align intended outcomes with the SX dimensions (Educational, Social, Personal)	- Clarify the boundaries of the evaluation effort - Ensure that time frames and responsibilities are well-defined - Connect each SX dimension to measurable objectives

(*continued*)

Table 1. (*continued*)

Phase/Step	Key Activities	Goals
1. Preliminary Planning (Confirm Resources & Stakeholders)	- Identify and brief evaluators, ensuring they have or acquire relevant expertise (e.g., heuristic evaluation, group walkthroughs, perspective-based inspection, interviews, surveys, focus groups) - Confirm availability of facilities, technological tools, and support staff	- Make certain the team has the necessary skill sets and resources to execute the evaluation - Plan for any required training or onboarding of evaluators - Secure logistical and technological support for a smooth evaluation process
2. Inspection Phase (Heuristic Evaluation)	**- Prepare Relevant Heuristics**: Compile usability and educational heuristics; include social interaction and personal growth considerations; account for cultural differences in usability, feedback, and communication **- Conduct Individual Evaluations:** Have each evaluator independently review the platform, categorizing findings under Educational, Social, Personal dimensions and noting cultural implications **- Aggregate & Prioritize Results:** Consolidate evaluators' reports, categorize issues by severity, and highlight any culturally sensitive concerns	- Identify usability strengths/weaknesses prior to involving students - Maintain focus on the three SX dimensions to ensure comprehensive coverage - Recognize potential cultural misalignments early (e.g., language nuances, culturally sensitive content)
2. Inspection Phase (Group-Based Expert Walkthrough)	- **Plan the Walkthrough**: Invite experts with diverse backgrounds and cross-cultural perspectives if possible; set clear objectives (e.g., customer journey from sign-in through learning, social interaction, and personal engagement) - **Facilitate Review**: Guide participants through task scenarios aligned with Educational, Social, Personal dimensions, discussing cultural norms (e.g., uncertainty avoidance, power distance) - **Document Observations**: Record suggestions focusing on clarity, inclusivity, motivation; label cultural variability as needed	- Gain a deeper, collective understanding of how the system performs under expert scrutiny - Leverage group dynamics to refine or validate individual findings from heuristic evaluation - Integrate multiple cultural viewpoints to address potential biases in design or functionality

(*continued*)

Table 1. (*continued*)

Phase/Step	Key Activities	Goals
2. Inspection Phase (Perspective-Based Inspection)	- **Identify Relevant Perspectives**: At minimum, include Instructor's/Educational, Peer/Social, Individual/Personal viewpoints; incorporate Hofstede's cultural dimensions into each perspective - **Assign Reviewers**: Select evaluators who can effectively represent each viewpoint; brief them on engagement, inclusivity, cultural respect - **Synthesize Findings**: Compile feedback, organize by perspective, and highlight overlapping or contradictory issues; note recurring cultural considerations	- Ensure a multifaceted appraisal of the platform or system - Capture the interplay between pedagogical objectives, social collaboration, and individual learner autonomy - Address cultural expectations across all perspectives to avoid narrow or incomplete recommendations
3. Student Testing Phase (Interviews)	- **Plan Interview Guidelines**: Develop open-ended questions on academic needs, sense of community, personal well-being, cultural comfort levels, communication preferences, and support networks - **Conduct & Record Interviews**: Schedule small-group or individual sessions; obtain informed consent; document cultural/contextual cues - **Summarize Findings**: Perform initial thematic analysis, noting recurring strengths/weaknesses and cultural sensitivities; categorize them under Educational, Social, Personal dimensions	- Elicit in-depth qualitative insights that may not emerge through surveys alone - Uncover specific cultural nuances and personal motivations influencing student experience - Map interview themes to the SX dimensions for later triangulation with other data

(*continued*)

Table 1. (*continued*)

Phase/Step	Key Activities	Goals
3. Student Testing Phase (Surveys)	- **Design Survey Instrument**: Use quantitative scales (e.g., Likert) with optional comment fields; ensure wording is culturally and linguistically accessible. It is essential to use specialized tools for assessing the Student Experience (SX), such as the SX evaluation survey incorporating cultural aspects that we are currently developing - **Distribute & Collect**: Provide clear instructions about purpose and duration; encourage broad participation to cover various cultural and demographic segments - **Analyze & Interpret**: Calculate descriptive statistics, identify patterns, and compare with inspection-phase findings; note correlations with Hofstede's dimensions	- Gather large-scale, representative feedback on user satisfaction and perceived issues - Confirm or challenge expert evaluations using quantitative data - Pinpoint how cultural factors (e.g., power distance) correlate with student satisfaction in Educational, Social, and Personal domains
3. Student Testing Phase (Focus Group)	- **Assemble Diverse Groups**: Aim for variety in academic background, cultural heritage, and platform experience; appoint a facilitator who can encourage equitable participation across cultural norms - **Moderate Sessions**: Start with broad questions related to Educational, Social, and Personal domains; probe specific features (User Interfaces (UI), collaboration, motivation) - **Capture Dynamics & Consensus**: Observe dominant voices or cultural communication patterns; document conclusions and debates, noting Hofstede's dimensions	- Engage students in collective dialogue to refine understanding of the platform's effectiveness and inclusivity - Validate or supplement findings from interviews and surveys with real-time interaction and discussion - Identify how different cultural values shape learning styles, motivation, and communication

(*continued*)

Table 1. (*continued*)

Phase/Step	Key Activities	Goals
4. Integration of Qualitative & Quantitative Data	- **Triangulate Inspection & Student Feedback**: Compare expert findings with interview, survey, and focus group results; note areas of agreement or discrepancy - **Organize Data by SX Dimensions**: Create clusters for Educational, Social, Personal, integrating cultural insights drawn from Hofstede's model - **Generate Actionable Recommendations**: Develop evidence-based proposals for improvements, tagging each recommendation by relevant SX dimension and cultural consideration - **Report Outcomes & Next Steps**: Present findings to stakeholders; highlight priorities, propose timelines, and assign responsibilities	- Achieve a comprehensive view by synthesizing both expert-driven and student-driven data - Confirm which issues are universally significant and which are culturally specific - Provide clear, prioritized action items to enhance the overall student experience (Educational, Social, Personal) while respecting cultural diversity

5.2 Considerations for Future Evaluation and Validation

In planning our upcoming —and iterative— evaluation and validation process for the checklist, we have carefully incorporated the criteria outlined in the Western Michigan University Checklist Project [24]. By grounding our methodology in these established standards, we seek to enhance the rigor and reliability of our approach. Table 2 offers a detailed overview of the specific measures we have —or will have— implemented for each criterion, thereby ensuring alignment with recognized best practices and facilitating a clear, comprehensive assessment.

It is paramount to emphasize that our commitment to inclusivity in research and higher education extends beyond merely acknowledging the diverse composition of our student body and the cultural dimensions of the methodology underpinning our proposal. We believe it is essential to involve evaluators of varied genders and cultural backgrounds—domestic and international—to ensure a comprehensive, equitable perspective. This approach aligns with the guidelines promoted by international organizations dedicated to improving the quality of education [28] and advancing the organic development of communities.

Furthermore, we wish to underscore that the methodological framework incorporated into this checklist directly addresses three of UNESCO's Sustainable Development Goals: Quality Education, Gender Equality, and Reduced Inequalities. These objectives are commonly factored into the accreditation processes for HEIs and, in our view, should

be explicitly considered when assessing our tool. By making these considerations transparent, we reinforce the value of our proposal and support the broader aim of fostering a more inclusive and effective educational environment.

Table 2. Considerations for the checklist evaluation and validation [24].

Criteria	Consideration
Appropriateness of evaluation content	We formulate our checklist so that it targets at least one distinct evaluation objective or an overarching activity that spans multiple tasks/activities. We will simplify intricate subject matter and highlight essential points, ensuring that the checklist is both user-friendly and grounded in credible sources, including our direct professional experiences. Throughout this process, we will remain consistent with the program evaluation standards of Yarbrough et al. [25] as well as the American Evaluation Association's Guiding Principles for Evaluators [26] and Statement on Cultural Competence in Evaluation [27]. In doing so, we commit to avoiding any undue emphasis on a single evaluative methodology, unless our intention is expressly to facilitate the use of that particular approach
Clarity of purpose	We propose the following measures to ensure our checklist aligns with the specified validation requirements. First, we will design a concise heading that unambiguously conveys the checklist's focus, fulfilling the guideline that "a succinct title clearly identifies what the checklist is about." Next, we will provide a short introductory section that explains its purpose, detailing both the situations in which it is applicable and the correct method of deployment. This introduction will also include cautions regarding improper usage, as suggested in the reference that it should clarify "how it should be used (including caveats about how it should not be used if needed)." In addition, we will identify the target audience so users understand who should employ the checklist. By combining a precise title, a brief contextual overview, and a defined user scope, we will create a tool that is both straightforward and fit for its intended purpose
Completeness and relevance	We propose a systematic approach that begins by thoroughly identifying the scope of the tasks we aim to evaluate, ensuring "All essential aspects of the evaluation tasks are addressed." We then refine the checklist details so that "All content is pertinent to what users need to do to complete the tasks," thus maintaining relevance and clarity. Throughout development, we conduct iterative reviews, consult with key stakeholders for feedback, and incorporate any necessary adjustments to guarantee that each step of the process remains purposeful. By actively involving potential users, we verify the practicality of every item in the checklist, confirm that directions are free of ambiguity, and validate that the outlined steps directly correspond to real-world needs. This collaborative and reflective methodology allows us to create a robust checklist that aligns with the stated references while supporting reliable and consistent performance across all intended activities

(*continued*)

Table 2. (*continued*)

Criteria	Consideration
Organization	We will arrange our checklist so that its substance is conveyed in a coherent progression, whether by overall concept or by chronological stages. We will also segment the material into distinct parts identified by succinct yet descriptive headers, which will facilitate swift navigation and comprehension. Furthermore, we will dissect intricate activities or elements into multiple simpler segments, ensuring that no single component overwhelms the user. By adhering to these measures, we confirm that our checklist meets the requirements for logical structuring, clear organization, and manageable complexity
Clarity of writing	In our efforts to design the checklist, we will concentrate on issuing clear directives for users rather than posing reflective inquiries. We will embrace plain language rather than technical terminology, while ensuring our verbs convey immediate, action-oriented tasks. We will also be deliberate in choosing precise words and applying them consistently, offering concise explanations for any concepts that may not be self-explanatory. Furthermore, we will maintain brevity and clarity in our sentence structure, so that the content remains accessible and straightforward for all readers
References and sources	In developing our checklist, we ensure that every reference consulted during its creation is properly acknowledged. We will also include a curated set of supplementary materials for anyone seeking to delve more deeply into the subject area. Moreover, we intend to provide a recommended way to cite the checklist—either at the outset or at the conclusion—to maintain consistency in scholarly usage. Finally, we will make our own contact details readily accessible so that readers can direct questions or feedback to us without difficulty

6 Study Limitations

The checklist developed in this study and the broader methodology in which it is embedded is specifically tailored to the SX of undergraduate learners. Consequently, the theoretical foundations informing this approach may render its application outside the undergraduate context susceptible to imprecise or limited results. Furthermore, the current work represents a preliminary version of the checklist, falling within the early stages of a comprehensive development process. As it has yet to undergo full review, refinement, and validation, the checklist remains subject to potential revisions before reaching its final form.

7 Conclusion and Future Work

In the CDC's first five stages (I–V), we focused on the checklist task, generated a preliminary list of checkpoints, classified and refined them according to key criteria, and established a logical order based on the evaluation methodology flow. Building on

these foundations, the following stages (VI–XII) will involve engaging in initial reviews by potential users and experts, revising and formatting the checklist to meet specific practical needs, conducting a thorough evaluation, and finalizing the instrument. Once finalized (Stage X), the checklist will be disseminated (Stage XI) so that it may be employed in diverse educational settings, with mechanisms in place for periodic review and revision (Stage XII) to sustain its long-term relevance and efficacy.

This study presents a checklist developed upon a robust theoretical and conceptual foundation embedded within a holistic methodology for evaluating the student experience. The design process underwent meticulous review and assessment by subject-matter experts representing diverse gender configurations, professional areas, academic fields, and cultural contexts. The proposed tool aims to ensure reliability, versatility, and broad applicability across multiple educational settings by adhering to internationally recognized guidelines. Collectively, these measures reinforce the checklist's capacity to provide comprehensive insights into the student experience, thereby serving as a valuable resource for educators, administrators, and researchers seeking to enhance learning environments.

Moving forward, our primary objective is to apply the checklist in authentic educational contexts to test its efficacy under real-world conditions. Such field-based deployment will allow us to gather meaningful data and assess the tool's performance and impact on student outcomes. Additionally, this implementation phase will provide critical insights into its adaptability across diverse academic and cultural settings, ensuring its relevance in varied learning environments. In conjunction with these analyses, we plan to refine the checklist iteratively by revisiting and adjusting individual items as necessary. To maintain methodological rigor, the evaluation process will incorporate a structured approach that accounts for culture-related variables [29], reinforcing the checklist's validity in comprehensively assessing the student experience. This cyclical process of application and improvement will help maximize the reliability and practical utility of the method, ensuring that it remains responsive to evolving educational challenges and maintains a consistently high standard of quality.

Acknowledgments. The authors of this article wish to extend their sincere gratitude to all the professionals, faculty members, and students, as well as the institutions to which they belong, for their invaluable collaboration in developing the methodology that underpins this checklist. We also wish to acknowledge the Escuela de Ingeniería Informática at the Pontificia Universidad Católica de Valparaíso for providing the financial support that made this work possible.

References

1. Parasuraman, A., Zeithaml, V.A., Berry, L.L.: A conceptual model of service quality and its implications for future research. J. Mark. **49**(4), 41–50 (1985)
2. Parasuraman, A., Zeithaml, V.A., Berry, L.L.: Servqual: a multiple-item scale measuring consumer perceptions of service quality. J. Retail. **64**, 12–40 (1988)
3. ISO 9241-210: Ergonomics of human-system interaction—Part 11: Usability: Definitions and concepts. International Organization for Standardization, Geneva (2018)
4. Nielsen Norman Group: How channels, devices, and touchpoints impact the customer journey. https://www.nngroup.com/articles/channels-devices-touchpoints/. Accessed 21 Jan 2025

5. Hill, N., Roche, G., Allen, L.: Customer Satisfaction. The Customer Experience Through the Customer's Eye, 1st edn. Cogent Publishing, Abingdon (2007)
6. Gentile, C., Spiller, N., Noci, G.: How to sustain the customer experience: an overview of experience components that co-create value with the customer. Eur. Manag. J. **25**, 395–410 (2007)
7. Sujata, J.: Customer experience management: an exploratory study on the parameters affecting customer experience for cellular mobile services of a telecom company. Procedia Soc. Behav. Sci. **133**, 392–399 (2014)
8. Harvey, L., Knight, P.T.: Transforming Higher Education. Society for Research into Higher Education/Open University Press, London (1996)
9. Arambewela, R., Maringe, F.: Mind the gap: Staff and postgraduate perceptions of student experience in higher education. High. Educ. Rev. **44**, 63–83 (2012)
10. Matus, N., Rusu, C., Cano, S.: Student eXperience: a systematic literature review. Appl. Sci. **11**(20), 9543 (2021)
11. UNESCO: The Contribution of Higher Education to the SDGs. https://www.iesalc.unesco.org/en/sustainable-development/contribution-higher-education-sdgs. Accessed 21 Jan 2025
12. UNESCO: The 17 Goals: Sustainable Development. https://sdgs.un.org/goals. Accessed 21 Jan 2025
13. Hofstede, G., Hofstede, G.J., Minkov, M.: Cultures and Organizations. Software of the Mind. McGraw Hill (2010)
14. Hofstede, G.: Culture's Consequences: Comparing Values, Behaviours, Institutions, and Organizations Across Nations, 2nd edn. Sage, London (1980)
15. Taras, V., Rowney, J., Steel, P.: Half a century of measuring culture: review of approaches, challenges, and limitations based on the analysis of 121 instruments for quantifying culture. J. Int. Manag. **15**, 357–373 (2009)
16. The Culture Factor: Frequently asked questions. https://www.theculturefactor.com/frequently-asked-questions. Accessed 21 Jan 2025
17. Hofstede, G.: Masculinity and Femininity: The Taboo Dimension of National Cultures. Sage, London (1998)
18. Matus, N., Rusu, C., Botella, F.: Towards well-being and inclusion in the educational system: a preliminary methodology for evaluating student eXperience considering cultural factors. In: Coman, A., Vasilache, S. (eds.) Social Computing and Social Media. HCII 2023. Lecture Notes in Computer Science, vol. 14026, pp. 103–115. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-35927-9_8
19. Matus, N., Rusu, C., Botella, F.: Proposing a SX model with cultural factors. Appl. Sci. **13**, 3713 (2023). https://doi.org/10.3390/app13063713
20. Matus, N., Ito, A., Rusu, C.: Analyzing the impact of culture on students: towards a student eXperience holistic model. In: Meiselwitz, G. (ed.) Social Computing and Social Media: Applications in Education and Commerce. HCII 2022. Lecture Notes in Computer Science, vol. 13316, pp. 127–135. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-05064-0_10
21. Almeida, R.L.A., Andrade, R.M.C., Darin, T.G.R., Paiva, J.O.V.: CHASE: checklist to assess user experience in IoT environments. In: New Ideas and Emerging Results (ICSE-NIER 2020), pp. 41–44. ACM, New York (2020). https://doi.org/10.1145/3377816.3381737
22. Richardson, C., Campbell-Yeo, M., Smit, M.: Mobile application user experience checklist: a tool to assess attention to core UX principles. Int. J. Hum. Comput. Interact. **37**(13), 1283–1290 (2021)
23. Stufflebeam, D.L.: Guidelines for developing evaluation checklists: the checklists development checklist (CDC). Western Michigan University (2000). https://files.wmich.edu/s3fs-public/attachments/u350/2014/guidelines_cdc.pdf. Accessed 21 Jan 2025

24. Western Michigan University: Checklist validation process. https://wmich.edu/evaluation/checklists/checklistsvalidation. Accessed 21 Jan 2025
25. Yarbrough, D.B., Shulha, L.M., Hopson, R.K., Caruthers, F.A.: The Program Evaluation Standards: A Guide for Evaluators and Evaluation Users, 3rd edn. Sage, Thousand Oaks (2011)
26. American Evaluation Association: Guiding principles for evaluators (2013). https://www.eval.org/About/Guiding-Principles. Accessed 21 Jan 2025
27. American Evaluation Association: Statement on cultural competence in evaluation (2011). https://www.eval.org/Community/Volunteer/Statement-on-Cultural-Competence-in-Evaluation. Accessed 21 January 2025
28. European Commission: A plan for achieving gender equality in the research and innovation system. https://projects.research-and-innovation.ec.europa.eu/en/projects/success-stories/all/plan-achieving-gender-equality-research-and-innovation-system. Accessed 21 Jan 2025
29. Frierson, H., Hood, S., Hughes, G.: A guide to conducting culturally responsive evaluation. In: Frechtling, J. (ed.) The 2002 User-Friendly Handbook for Project Evaluation, pp. 63–73. National Science Foundation (2002)

Human-Robot Interaction in Higher Education: A Literature Review

Sebastian Matus(✉) and Sandra Cano

Pontificia Universidad Catolica de Valparaiso, 2340000 Valparaiso, Chile
sebastian.matus@pucv.cl

Abstract. This review explores the growing role of robots in higher education, emphasizing their potential to improve learning outcomes and transform teaching methodologies. Analyzing existing research, this synthesis explores the diverse roles of robots in higher education: intelligent tutors, collaborative partners, and social companions. Additionally, the paper identifies the critical elements necessary for successful integration, including interaction design ethical considerations, the role of educators, and technical issues. By analyzing current findings and addressing gaps in the literature, this study offers a framework for deploying social robots in higher education, providing guidelines for educators and researchers to ensure that these technologies contribute meaningfully to the learning experience. The review concludes with recommendations for future research and best practices in the application of robots in university classrooms.

Keyword: Human-Robot Interaction · Higher Education · Social Robots

1 Introduction

Integrating robots and other intelligent systems into the higher education landscape has been a topic of growing interest and research in recent years. Studies have explored the potential of these technologies to enhance teaching and learning, particularly in STEM disciplines [1].

One key area of focus within this field is the study of human-robot interaction in higher education. Human-robot interaction aims to leverage the complementary strengths of humans and robots. Robots can assist with tasks requiring precision, speed, and force, while humans contribute their experience, knowledge, intuition, and adaptability [2].

Social robots, defined as autonomous or semi-autonomous entities capable of interacting with humans while adhering to social norms [3], have become increasingly prevalent in educational settings. Their applications span tutoring, collaboration, emotional support, and adaptive learning, significantly transforming traditional education. Robots like NAO and Pepper have been extensively employed as tutors, capable of delivering lectures, providing personalized feedback, and adapting content to meet individual learning needs [4, 5]. These robots leverage real-time data to adjust teaching strategies, dynamically improving student comprehension and retention [6]. Additionally, social robots enhance group dynamics by mediating and moderating collaborative activities,

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 236–256, 2025.
https://doi.org/10.1007/978-3-031-93539-8_16

leading to improved time management, inclusivity, and teamwork outcomes [7]. For example, robots have effectively facilitated complex problem-solving tasks, such as Planning Poker sessions for software engineering students [6].

Equipped with emotion recognition technologies, robots can detect stress, anxiety, and disengagement, providing tailored interventions such as mindfulness exercises or empathetic responses [8]. Therapeutic robots like Paro and KiliRo have successfully reduced student anxiety, creating a supportive learning environment [9]. These technological advancements have also had profound educational impacts, addressing traditional challenges and fostering more inclusive and adaptive learning environments. Robots significantly enhance learning outcomes by delivering consistent, tailored instruction and providing real-time feedback [10]. Gamified robotic platforms make abstract concepts more accessible and engaging, particularly in STEM disciplines [11]. Furthermore, multimodal interaction capabilities, including gestures, speech, and tactile feedback, increase student engagement and maintain attention [12]. Integrating gamification elements into robotic interactions has fostered greater motivation and persistence in learning activities [6].

Robots tailored for multilingual environments and culturally specific non-verbal communication styles increase their relevance and acceptance in diverse educational settings. For example, culturally responsive robots have been found to boost participation and engagement in English Language Learners (ELL) programs [13]. Advances in robotics and AI have significantly improved the functionality of social robots in educational contexts. AI-driven systems allow robots to personalize teaching based on performance metrics, fostering individualized learning experiences [14]. Advanced machine learning algorithms identify knowledge gaps and provide targeted feedback [5]. Emotion recognition technologies enable robots to analyze facial expressions, tone, and physiological signals, allowing them to gauge student emotions and offer empathetic responses. This creates emotionally intelligent robots capable of real-time emotional support [15]. Robots equipped with tactile, visual, and auditory modalities create engaging user experiences that maintain interest and participation [16]. Advances in tactile sensors enable robots to respond to physical interactions, enhancing safety and communication [17]. Gamification features, such as leaderboards and rewards, integrated into robotic systems have proven effective in enhancing motivation, particularly in STEM education [18].

Adopting social robots in higher education requires addressing significant ethical and cultural challenges. Privacy concerns remain a barrier to widespread adoption. Robots designed with privacy-by-design principles, such as anonymized data collection and secure storage, are more likely to gain acceptance [19]. Integrating culturally specific gestures, languages, and teaching styles into robots increases trust and usability among diverse student populations [20]. Addressing algorithmic bias, autonomy, and overreliance on robots necessitates clear guidelines to ensure responsible AI and robotic use. Despite their potential, social robots face several challenges in education. Current robots often lack advanced contextual understanding and emotional depth, which limits their ability to engage students fully. Furthermore, developing and deploying social robots remain expensive, posing challenges for under-resourced institutions. Affordable and scalable robotic platforms are needed to address this issue.

However, while their educational impacts and technological capabilities offer tremendous promise, addressing ethical, cultural, and technical challenges is critical for widespread adoption [21]. Future advancements in affordability, emotional intelligence, and inclusivity will determine the long-term success of social robots in transforming education.

This article is organized as follows: Sect. 2 presents the theoretical background, Sect. 3 describes the Objectives and Research Questions, Sect. 4 describes the methodology, and Sect. 5 presents the results obtained. Section 6 presents a discussion, and finally, Sect. 7 highlights the conclusions.

2 Theoretical Background

Human-robot interaction (HRI) is a dynamic and interdisciplinary field that focuses on designing, implementing, and studying interactions between humans and robots. With robots increasingly present in industries, healthcare, and education, HRI has become central to ensuring these systems are intuitive, effective, and human-centric. The relevance of HRI lies in its ability to bridge technological advances with human needs, fostering trust and usability while addressing ethical considerations [22, 23].

HRI research has progressed from functional, task-oriented interactions to emotionally and socially intelligent systems that can adapt to human behaviors and environments. This evolution is pivotal as robots move beyond industrial applications into domains requiring nuanced interpersonal exchanges, such as education and healthcare [24, 25].

In education, HRI offers immense potential to enhance teaching and learning experiences through interactive, adaptive, and engaging technologies. Robots can act as tutors, learning companions, and facilitators, complementing traditional teaching methods while providing personalized instruction [21, 26]. Their ability to deliver consistent feedback, adapt to individual learning styles, and foster engagement makes them particularly effective in supporting diverse educational needs.

Several studies from the uploaded articles illustrate these benefits. Enhanced Engagement and Motivation: Robots integrated into classroom activities, such as language learning and collaborative tasks, have increased student motivation and participation.

Skill Development: Robots are particularly effective in developing cognitive, communication, and presentation skills. Studies have demonstrated improved oral presentation performance and language proficiency through robot-assisted learning.

Interactive Learning Environments: Gamified and narrative-driven educational approaches using robots have proven successful in sustaining long-term engagement and fostering active learning.

3 Objectives

The objective of this systematic literature review is to examine the research that has been carried out on human-robot interaction used for higher education. The integration of social robots in higher education presents unique opportunities and challenges. The following research questions aim to explore the multifaceted roles of robots in academic settings, focusing on their impact on learning and motivation, the definition of their roles, and the critical elements required for successful integration. Our Research Questions are:

1. Does the Interaction with a Robot Support Students' Learning and Motivation?
2. What Should the Robot's Role Be When Integrated into Higher Education?
3. What Are the Key Elements to Consider in Robot Integration in Higher Education?

4 Methodology

The literature was systematically reviewed using PRISMA methodology [27]. SCOPUS and WoS were the databases used. The inclusion and exclusion criteria (Table 1 and Table 2, respectively), selected information sources (Subsect. 4.2), search strategy, and selection process are specified for replicability.

4.1 Study Selection Criteria

We included the conditions listed in Table 1 to perform the systematic literature review under the PRISMA methodology and develop a general knowledge of the concepts we were working with the types of papers presented in Table 2 were excluded.

Table 1. Inclusion criteria.

ID	Inclusion criteria
IN1	Articles focused on robots in higher education or related educational settings
IN2	Studies addressing robot roles as tutors, collaborators, or emotional support
IN3	Research on adaptive learning, gamification, or emotion recognition
IN4	Articles published in journals
IN5	Empirical studies, reviews, or case studies
IN6	Articles published in the last 5 year

Table 2. Exclusion criteria.

ID	Inclusion criteria
EX1	Articles not directly related to education or higher education
EX2	Studies focused solely on industrial robots or robotics unrelated to social interactions
EX3	Duplicate articles or redundant data

4.2 Data Sources

To conduct this systematic literature review, we searched for scientific papers on Scopus and Web of Science databases. For sources, we considered all documents found using the search string. Additionally, we selected articles published during the last five years.

4.3 Search Strings

We formulated the search strings based on the relevant topics to our systematic literature review. We determined a set of specific keywords to use in our queries, i.e., "human-robot interaction," "higher education," "university students," "social robot," and "emotional robots" that would be useful to answer our research questions.

These strings were focused on finding studies that analyzed or experimented with a robot in a higher educational context. Table 3 presents the specific search strings used in the selected databases.

Table 3. Inclusion criteria.

Database	Inclusion criteria
Scopus	TITLE-ABS-KEY ("human-robot interaction" OR "social robot" OR "social robotics" OR "robot tutor" OR "intelligent agent" OR "robotic tutor" OR "assistive robot" OR "Emotional robots" OR "affective computing") AND ("higher education" OR "university students" OR "college students") AND PUBYEAR > 2019 AND PUBYEAR < 2026 AND (LIMIT-TO (DOCTYPE, "ar")) AND (LIMIT-TO (LANGUAGE, "English"))
WoS	("human-robot interaction" OR "social robot" OR "social robotics" OR "robot tutor" OR "intelligent agent" OR "robotic tutor" OR "assistive robot" OR "Emotional robots" OR "social interaction" OR "affective computing") AND ("higher education" OR "university students" OR "college students")

4.4 Article Selection

Including both databases (SCOPUS and WoS), 1306 articles were found, 824 in WoS and 513 in SCOPUS. The registers were downloaded and organized in an Excel sheet; duplicates were eliminated, and Inclusion-Exclusion criteria were applied by reading the abstract of the articles. After that, 105 articles were left; a double check was performed using the inclusion-exclusion criteria again, but the entire article was read this time. Finally, 51 articles were assessed for eligibility; the analysis used those same 51 (Fig. 1).

5 Results

Table 4 shows the analysis of the 51 articles; Column 1 shows the year, Column 2 shows the year, and Column 3 shows the study's objective. Column 4 shows the robot role in the study, grouped in the next categories:

- Tutor: a robot that takes the role of a lecturer, teacher's assistant, or helps the student inside a classroom.
- Social robot: a robot that takes the role of social support or a companion.
- Facilitator: a robot that supports the student in a task in a role different from the tutor (for example, cheer him up if he fails).

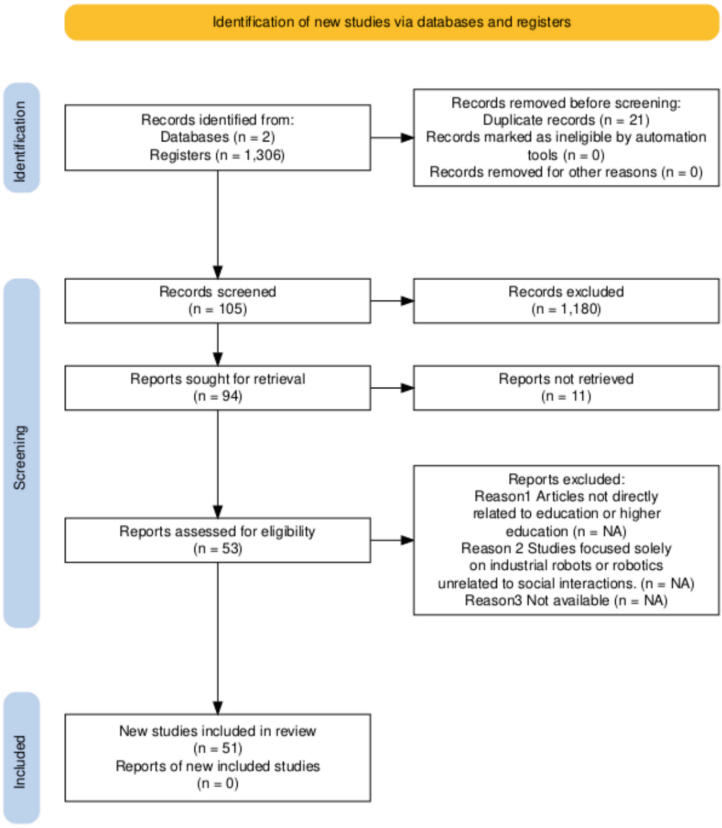


Fig. 1. Flow chart with the results of the article selection process.

- Educational tool: a robot that can be seen as a tool.
- Other: any role that doesn't fit in any previous category.

Column 5 shows whether the study includes (or follows) a learning theory (constructivism, social learning theory, connectivism learning theory, Self-Determination Theory, or other theory, not necessarily a learning one). Finally, column 6 shows the students who were part of the study: their course (or career), study grade, N, and age.

5.1 Answering the Research Questions

Question 1: Does the interaction with a robot support student's learning and motivation?

Integrating robots in education has sparked interest in their potential to enhance students' learning experiences and motivation. Social robots, equipped with artificial intelligence and interactive capabilities, are increasingly utilized in classrooms for tasks ranging from tutoring and assessment to fostering engagement. This report synthesizes findings from 51 research articles to evaluate whether interaction with robots supports students' learning and motivation, highlighting key evidence and practical implications.

Table 4. Article's analysis

ID	Year	Study's objective	Robot role	Learning theory included	Students
1	2024	Exploring how robotic tutors can enhance learning outcomes in software engineering education	Tutor	Not explicitly stated	Software engineering
2	2025	Analyze the use of entertainment robots with artificial intelligence	Tutor	Not explicitly stated	Not explicitly stated
3	2024	Investigating integrating musical education with virtual reality (VR)	Social robot	Not explicitly stated	Not explicitly stated
4	2024	Exploring the influence of robot mood expressions	Facilitator	Not explicitly stated	Not explicitly stated
5	2024	Investigate how nonverbal feedback from a humanoid robot influences human behavior and performance	Other	Other	Not explicitly stated
6	2024	Introduce and evaluate a humanoid head platform designed for robotics teaching and research	Educational Tool	Other	Undergraduate and master's students in robotics
7	2024	Explore the potential of telepresence robots in libraries	Educational Tool	Not explicitly mentioned	Library and information science
8	2024	Examine the integration of a social robot and gamification in adult learning environment	Facilitator	Other	Not explicitly stated
9	2024	To investigate the use of telepresence robots in classrooms	Educational Tool	Social constructivist theory	22 students
10	2024	To analyze interactional practices in robot-mediated hybrid language classrooms	Learning Companion	Other	Foreign language students

(*continued*)

Table 4. (*continued*)

ID	Year	Study's objective	Robot role	Learning theory included	Students
11	2023	Investigate how a robot's charismatic versus non-charismatic speaking affects creativity and performance	Facilitator	Social learning theory	100 students from multiple university programs across Denmark and Switzerland
12	2023	To investigate the relationship between human-robot interaction (HRI) and sustainability	Social robot	Other	77 university students from Hungary
13	2023	To investigate the impact of visible cabling and storytelling content general perception of social robots	Social robot	Not explicitly stated	200 university students from the University of Würzburg and the Technical University of Berlin
14	2023	To evaluate Chinese higher education students' intention to adopt AI-based robots for educational purposes	Educational Tool	Not explicitly stated	348 university students from engineering and information technology fields in Chinese universities
15	2023	To explore how artificial intelligence (AI) and social robotics (SR) can act as vectors for fostering complex thinking competencies	Educational Tool	Not explicitly stated	Not explicitly stated
16	2023	To explore university students' perceptions of robotics in the context of higher education and the Industry 4.0 era	Social robot and educational tool	Other	365 university students from a South Korean university
17	2021	To analyze the impact of using the ARCS motivational model combined with problem-based learning (PBL) on human interaction with robots	Educational Tool and Facilitator	Other	38 university students from Mechanical Engineering at a university in Taipei City, Taiwan

(*continued*)

Table 4. (*continued*)

ID	Year	Study's objective	Robot role	Learning theory included	Students
18	2022	To implement a Project-Based Learning (PBL) approach in an educational robotics lab for undergraduate students	Social robot and educational tool	The study employs Constructivist Learning Theory	The study involved 18 undergraduate students from an Electronic Engineering program
19	2022	To explore the effectiveness of a physical versus virtual robot invigilator in monitoring exam settings	Other	Other	76 participants in an online study and 26 participants in an in-person evaluation
20	2022	To explore the potential roles and applications of social drones in educational settings	Tutor	Other	20 participants, university students and graduates
21	2021	To evaluate the impact of robot-tutors versus human-tutors on knowledge acquisition, student enjoyment, and surprise levels	Tutor	Other	138 first-year university students, aged 18–28, enrolled in an educational department
22	2022	To evaluate the effectiveness of a robot-tutor compared to a human-tutor in enhancing student learning outcomes, enjoyment levels, and surprise levels	Tutor	Other	138 university freshmen from the Educational and Social Policies department
23	2022	To examine the effects of telepresence robots on student cooperation, comparing them with traditional videoconferencing tools	Facilitator	Other	122 university students from the Rotterdam University of Applied Sciences
24	2021	To explore adolescents' preferences for different embodiments of a social robot	Social robot	Other	66 adolescents aged 14–19 from three urban high schools and a university

(*continued*)

Table 4. (*continued*)

ID	Year	Study's objective	Robot role	Learning theory included	Students
25	2020	To systematically review the use of Socially Assistive Robots (SARs) in pre-tertiary education	Educational Tool	Constructivism and Constructionism	The systematic review included data from 1,333 participants across various studies
26	2020	To investigate the familiarity and interest of engineering students in social robotics	Social robot	Not explicit stated	72 Electrical Engineering students and 19 Computer Science students
27	2020	To develop and evaluate a robotic question support system designed to reduce students' hesitation in asking questions during lectures	Tutor	Constructivism	62 first- and second-year Japanese university students, with an average age of 19.5 years
28	2020	To examine the mediating role of robot anxiety in the relationship between social anxiety and attitudes toward interacting with robots	Social robot	Not explicit stated	286 university students from Istanbul, Turkey
29	2021	To assess the validity of the Negative Attitudes Toward Robots Scale (NARS) in a U.S. educational context	Other	Not explicit stated	54 undergraduate pre-service teachers
30	2020	To examine how students form interpersonal impressions of a social robot versus a human in the context of performance evaluations	Tutor	Other	58 undergraduate students with an average age of 21.65 years
31	2020	The study aims to understand how privacy concerns shape intentions to use social robots	Social robot	Not explicit stated	480 respondents

(*continued*)

Table 4. (*continued*)

ID	Year	Study's objective	Robot role	Learning theory included	Students
32	2020	To explore university students' preferences for educational robot design using a User-Centered Design (UCD) approach	Tutor	Other	116 university students, aged 17–40 years, from (STEM) fields, and sciences and humanities
33	2022	To introduce and evaluate a student engagement prediction model using Human-Robot Interaction (SEPM-HRI) in higher education	Facilitator	Not explicit stated	Not explicit stated
34	2024	To develop and evaluate an open-source Robotic Study Companion (RSC) designed for multimodal human-robot interaction	Facilitator	Social Learning	Not explicit stated
35	2022	To evaluate the effectiveness of adaptive robotic tutors in higher education	Tutor	Self-Determination Theory (SDT)	58 first-year undergraduate students
36	2020	To investigate the effectiveness of social robots as teaching assistants in higher education	Tutor	Other	36 university students
37	2023	To evaluate the effectiveness of a robotic positive psychology in improving students' psychological health	Social robot	Other	35 undergraduate students (age M = 18.94, SD = 1.43)
38	2024	To analyze the impact of using social robotics as a pedagogical tool to develop complex thinking skills	Facilitator	Other	51 graduate students

(*continued*)

Table 4. (*continued*)

ID	Year	Study's objective	Robot role	Learning theory included	Students
39	2023	To examine how privacy settings of social robots influence users' attitudes and willingness to self-disclose	Other	Other	320 students
40	2020	To investigate whether interacting with robot animal companions during final exams can help reduce stress	Social robot	Other	103 university students
41	2024	To explore university students' opinions and preferences regarding the use of intelligent agents	Social robot	Not	1,054 undergraduate students
42	2020	To evaluate the acceptance of humanoid robots in higher education	Tutor	Other	462 first-year university students
43	2021	To explore the use of robotics in physical therapy through a Project-Based Learning (PBL) approach	Other	Constructivist	Undergraduate engineering students from Tecnologico de Monterrey
44	2023	To investigate the effects of Robot-Supported Collaborative Learning (RSCL) in small groups,	Facilitator	Other	46 undergraduate university students, majority between 18 and 24 years old
45	2022	To investigate how robot autonomy influences ambivalent attitudes toward robots, examining how users simultaneously hold positive and negative evaluations about autonomous robots	Social robot	Other	The study involved 163 German university students

(*continued*)

Table 4. (*continued*)

ID	Year	Study's objective	Robot role	Learning theory included	Students
46	2024	To investigate the effectiveness of social robots as mental health coaches	Other	Other	36 university students
47	2021	To examine pre-service teachers' beliefs about the acceptability and adoption of social robots	Other	Other	121 first-year pre-service teachers students
48	2023	To examine how human participants interact with humanoid robots in a strategic decision-making setting	Other	Other	500 university students
49	2022	To develop an adaptive algorithm for humanoid robots (AHRs) in oral assessments	Facilitator	Other	Not explicit stated
50	2021	To examine the combined effects of social robots and gamification on motivation and engagement in higher education learning environments	Tutor	Self-Determination Theory (SDT)	80 university students enrolled in a Digital Media course
51	2024	To investigate the effects of robot-based multiple low-stakes assessments (Robot-MLSA) on students' oral presentation performance	Facilitator	Social Learning Theory	Not explicit stated

Key Findings. Improved Knowledge Retention: Several studies demonstrate that robot-assisted learning environments improve knowledge retention. For instance, the study on robot-based low-stakes assessments showed significant improvement in students' oral presentation skills compared to traditional computer-based methods [25].

Development of Communication Skills: Robots in language learning settings, such as Robot-Assisted Language Learning (RALL), significantly improved verbal and non-verbal communication skills. These effects were particularly pronounced when robots acted as conversational partners or facilitators [25, 28].

Real-Time Feedback: Robots that provide immediate, tailored feedback enhance students' understanding and application of complex concepts, such as those in STEM education and language learning [29–31].

Personalized Learning: Adaptive robot behavior, as highlighted in oral assessment scenarios, supports differentiated instruction, catering to student's unique needs and abilities [23].

Engagement Through Novelty: Robots in the classroom often increase initial engagement due to their novelty. Studies report heightened attention and interest, particularly in younger learners, when robots are integrated into learning tasks [32, 33].

Gamification and Interactive Elements: Combining robots with gamification elements further boosts motivation by making learning activities enjoyable and interactive. However, caution is needed, as overusing gamification can sometimes reduce long-term engagement [24, 34].

Reduction in Anxiety: Robots' non-judgmental nature helps reduce students' anxiety, especially in assessments or public speaking tasks, encouraging participation and self-expression [35].

Despite the previous findings, there are some limitations and also mixed findings:

- Sustained Motivation: While robots often increase short-term motivation, their novelty may wear off over time, requiring continuous innovation in instructional design to maintain engagement [22, 36].
- Cultural and Contextual Variability: Acceptance of robots and their impact on learning varies across cultural and educational contexts. Pre-service teachers, for example, expressed reluctance to adopt robots in classrooms due to perceived social and ethical concerns [31].
- Technical Challenges: Dependence on technology, such as internet reliability and robot functionality, can hinder seamless integration and learning effectiveness [25, 37].

Interaction with robots supports students' learning and motivation through enhanced engagement, personalized feedback, and reduced anxiety. However, to maximize these benefits, we should incorporate pedagogical principles, as robots should be integrated with well-designed instructional strategies that align with learning objectives. Focus on Sustainability: Innovations to sustain long-term motivation and engagement, such as periodic updates to robot functions and interaction modes, are essential. Address Ethical and Cultural Concerns: Educators should address potential biases and ethical implications to increase stakeholder acceptance. And finally, ensure Infrastructure Readiness: Reliable technological infrastructure is crucial for effective classroom robot deployment.

Question 2: What should the robot's role be when integrated into higher education? Robots in higher education have the potential to transform teaching and learning practices by acting as tutors, collaborative partners, and even social companions. Drawing on insights from 51 research articles, this report evaluates what roles robots should ideally play in higher education to maximize their impact while addressing pedagogical, ethical, and infrastructural considerations.

Key Roles for Robots in Higher Education. Tutor and Personalized Instructor: Robots can act as tutors, providing personalized Instruction. Adaptive algorithms allow robots to tailor learning materials and teaching methods to individual student needs, enhancing understanding and retention [23, 38, 39].

Assessment Facilitators: Robots can deliver formative and summative assessments while providing immediate, constructive feedback. For example, oral robot-based evaluations improved student performance and reduced anxiety [25]. Supplementary Educators: They can reinforce complex topics in STEM, language learning, and other disciplines through interactive sessions [30, 40].

Collaborative Learning Partner: Robots can foster collaboration by facilitating Group Work, acting as a mediator in group discussions, ensuring equitable participation, and guiding the group toward learning objectives [41, 42]. Encouraging Peer-to-Peer Interaction: Robots can simulate scenarios where students practice communication and critical thinking skills, enhancing their collaborative competencies [31, 35].

Social Companion: In higher education, robots can serve as social companions, providing emotional support by addressing student anxiety, especially in language learning and public speaking, by acting as a non-judgmental partner [22, 33]. Motivational Engagement: Robots' interactive capabilities can keep students motivated and engaged during long or challenging courses [32, 36].

Facilitator of Lifelong Learning. Robots can encourage lifelong learning by promoting self-directed learning, acting as a resource for students to explore topics independently, and being supported by robots' ability to answer queries and provide tailored resources [43, 44]. Enhancing Digital Literacy: Robots in higher education can prepare students for future technological demands by familiarizing them with AI systems and their applications [10, 34].

Research and Experimentation Tool. Robots can support research activities in higher education: Human-Robot Interaction Studies: Providing data and platforms for experiments in fields like psychology, engineering, and education [45, 46]. Innovative Applications: Encouraging interdisciplinary research by acting as a tool for developing novel applications in healthcare, robotics, and computational sciences [28, 47].

On the other hand, Robots must align with educational goals rather than serve as a technological novelty. Their role should complement, not replace, human instructors [30, 37]. There are ethical and social Implications, like acceptance; faculty and student perceptions of robots can vary, necessitating training and clear communication about their purpose [24, 31]. Bias and Inclusivity: Robots should be designed to avoid reinforcing biases and should support diverse learners effectively [23, 48].

Also, there are technical and infrastructural Challenges, like the integration with Existing Systems, as robots should seamlessly integrate with Learning Management Systems (LMS) and other educational tools [19, 29].

Cost and Maintenance: The financial and logistical aspects of deploying and maintaining robots must be carefully considered [25, 37].

Finally, there are some recommendations. First, to optimize their role in higher education, robots should act as Supplemental Tools, supporting rather than replacing human instructors by providing personalized and scalable learning opportunities. Also, they should facilitate active and collaborative learning, encouraging peer interaction and fostering skills like critical thinking and teamwork.

They must adapt to individual needs and utilize AI to provide tailored feedback, resources, and assessments; promote Ethical Integration by addressing cultural and ethical concerns to enhance acceptance and inclusivity. Support Research and Innovation, as they can serve as a platform for interdisciplinary research and innovation in various domains. Robots in higher education hold significant potential but require thoughtful integration to ensure they align with pedagogical objectives and enhance learning experiences.

Question 3: What are the key elements to consider in robot integration in higher education?

Integrating robots into higher education presents exciting opportunities to enhance teaching, learning, and research. However, successful integration requires careful planning and consideration of multiple elements. This report synthesizes findings from 51 research articles to outline the critical factors for effective robot integration in higher education settings.

Key Elements of Robot Integration. Pedagogical Design, as the alignment with Learning Objectives, needs to be precise; robots should complement the curriculum and support clearly defined learning outcomes, such as fostering critical thinking, collaboration, or knowledge retention [29, 30, 38]. Role definition is also vital; the robot's specific role—tutor, collaborator, or assessment facilitator—must be explicitly defined and aligned with pedagogical goals [23, 33, 39]. Finally, integration with Teaching Strategies, as robots should be incorporated into active learning approaches, such as problem-based learning or flipped classrooms, rather than as standalone tools [34, 40].

Student Engagement and Experience, like motivation and interactivity. Robots should be designed to enhance engagement through interactive features, including gamification and real-time feedback [24, 32, 35]. Reducing Anxiety: Robots can play a key role in creating a non-judgmental learning environment, particularly in language learning and oral assessments [28, 36]. Accessibility and Inclusivity: Robots should accommodate diverse student needs and ensure equitable access for all learners, including those with disabilities [37, 44].

There's technical and infrastructure readiness, like the integration with Existing Systems; robots should seamlessly integrate with Learning Management Systems (LMS) and other digital tools used in higher education [19, 29]. Reliability and Scalability: Robots must be reliable and scalable to accommodate varying class sizes and institutional needs [31, 37]. User-Friendly Interfaces: Both students and instructors should find robot interfaces intuitive and easy to use, minimizing the learning curve [10, 30].

As for the Faculty Training and Acceptance: Professional Development: Faculty must receive comprehensive training on using robots effectively in the classroom [22, 31].Addressing Resistance: Educators' concerns about robots replacing human instructors must be addressed through clear communication of their complementary role [39, 48]. Collaborative Design: Faculty should be involved in designing and deploying robotic systems to ensure alignment with teaching practices [41, 45].

Regardless with the ethical and Social Considerations, there is the Privacy and Data Security, as Institutions must ensure compliance with data privacy regulations, especially when robots collect and store student data [47, 49]. Bias Mitigation: Robots should be

designed to avoid reinforcing stereotypes or biases, promoting inclusivity and fairness [23, 48]. Cultural Sensitivity: Robots must be tailored to the cultural and linguistic context of the institution and its learners [33, 44].

In Assessment and Evaluation Measuring Effectiveness, Institutions must establish metrics to evaluate robots' impact on learning outcomes and student satisfaction [25, 30]. Continuous Feedback: Regular feedback from students and faculty should inform iterative improvements to robot design and functionality [32, 34]. Cost-Benefit Analysis: A thorough evaluation of costs versus benefits is essential to justify the investment in robotic systems [31, 37].

Finally, there are some recommendations: Successful robot integration in higher education requires careful attention to pedagogical alignment, student engagement, technical readiness, faculty training, ethical considerations, and ongoing assessment. By addressing these key elements, institutions can harness robots' potential to enhance teaching and learning while minimizing challenges and maximizing benefits.

This analysis provides a comprehensive overview of the elements critical to integrating robots effectively in higher education environments.

6 Discussion

While integrating robots into higher education holds significant promise, the 51 articles reviewed reveal several limitations and challenges warrant careful consideration. These insights can guide future research and implementation efforts to address potential pitfalls effectively.

Technological limitations: System Reliability: Robots may experience technical malfunctions, disrupting learning activities and reducing trust in their capabilities. Limited reliability can deter consistent use and lead to frustration among students and faculty [29, 37]. Integration with Existing Infrastructure: Many robots struggle to integrate seamlessly with Learning Management Systems (LMS) and other digital tools, hindering their adoption and usability [19, 30]. Maintenance and Updates: Regular maintenance and software updates are required to ensure functionality, posing logistical and financial burdens on institutions [10, 37].

Pedagogical challenges: Lack of Pedagogical Frameworks: The absence of standardized guidelines for incorporating robots into educational settings limits their effective deployment. Educators often lack clear strategies to align robots with curriculum objectives [34, 39].

Over-reliance on Robots: There is a risk of over-relying on robots, potentially diminishing the role of human instructors and affecting the quality of interpersonal interactions in the learning process [31, 48], and limited adaptability, some robots lack the flexibility to adapt to diverse student needs, such as accommodating different learning styles or cultural contexts [33, 44].

Ethical and Privacy Concerns: Data Privacy: Robots often collect and store sensitive student data, raising concerns about compliance with privacy regulations and the potential misuse of information [47, 49]. Student Autonomy: Overuse of robots may inadvertently undermine students' autonomy by creating dependency on robotic assistance for tasks they could otherwise perform independently [36, 40]. Cultural Sensitivity: Robots

designed without consideration for cultural differences may struggle to gain acceptance or fail to provide adequate support in diverse educational settings [33, 37].

Financial Constraints: High Initial Costs: Acquiring and deploying robots often require significant financial investment, which may not be feasible for underfunded institutions [10, 37]. Unequal Access: Institutions with limited resources may struggle to implement robotic solutions, exacerbating educational inequalities [25, 31].

Sustainability of Engagement: Novelty Effect: The initial excitement about robots often fades over time, requiring continuous updates to maintain student engagement [22, 36]. Scalability Issues: While robots may work well in small-scale trials, scaling their use to larger classrooms or entire institutions often reveals unforeseen challenges [31, 37].

7 Conclusions

Integrating robots into higher education represents a transformative opportunity to enhance teaching, learning, and research. Drawing from insights across 51 research articles, this review has highlighted both the significant potential and the challenges associated with robotic integration. The findings emphasize the importance of thoughtful, evidence-based approaches to ensure that robots are practical tools.

The successful integration of robots into higher education requires careful attention to several critical factors to ensure their effectiveness and sustainability. First, robots must be aligned with clearly defined learning objectives and pedagogical strategies, ensuring their use directly supports educational goals. Second, developers and institutions should prioritize ethical and inclusive design, addressing data privacy, cultural sensitivity, and inclusivity to foster acceptance and equitable access among diverse student populations. Third, comprehensive training programs and ongoing support for faculty and students are essential to overcoming resistance and effectively utilizing robotic technologies. Fourth, institutions should adopt sustainable models, exploring cost-effective and scalable solutions for deployment, such as partnerships with technology providers and open-source platforms. Lastly, continuous evaluation of robots' effectiveness and user satisfaction is necessary to guide iterative improvements. By addressing these factors, robots can be valuable in enhancing higher education.

Robots can potentially revolutionize higher education by enhancing the quality and accessibility of learning experiences. However, their integration must be approached with a balanced perspective, recognizing their strengths and challenges. By addressing the limitations and leveraging the opportunities identified in this review, institutions can create innovative, inclusive, and effective learning environments that prepare students for the demands of a rapidly evolving world.

References

1. Cruz-Ramírez, S.R., García-Martínez, M., Olais-Govea, J.M.: NAO robots as context to teach numerical methods. Int. J. Interact. Des. Manuf. IJIDeM **16**(4), 1337–1356 (2022). https://doi.org/10.1007/s12008-022-01065-y

2. Sharkawy, A.-N.: Human-Robot Interaction: Applications (2021). https://doi.org/10.48550/ARXIV.2102.00928
3. Bharatharaj, J., Sasthan Kutty, S.K., Munisamy, A., Krägeloh, C.U.: What do members of parliament in india think of robots? Validation of the Frankenstein Syndrome Questionnaire and comparison with other population groups. Int. J. Soc. Robot. **14**(9), 2009–2018 (2022). https://doi.org/10.1007/s12369-022-00921-x
4. Reyes, G.E.B., López, E., Ponce, P., Mazón, N.: Role assignment analysis of an assistive robotic platform in a high school mathematics class, through a gamification and usability evaluation. Int. J. Soc. Robot. **13**(5), 1063–1078 (2021). https://doi.org/10.1007/s12369-020-00698-x
5. Ponce, P., López-Orozco, C.F., Reyes, G.E.B., Lopez-Caudana, E., Parra, N.M., Molina, A.: Use of robotic platforms as a tool to support STEM and physical education in developed countries: a descriptive analysis. Sensors **22**(3), 1037 (2022). https://doi.org/10.3390/s22031037
6. Balalaieva, O., Mochan, T., Hryhorenko, T., Andreikova, I., Paltseva, V., Podkovyroff, N.: Innovative pedagogical technologies – the most important resource in modernizing the training of a modern specialist. Rev. Amazon Investiga **12**(63), 67–76 (2023). https://doi.org/10.34069/AI/2023.63.03.6
7. Buchem, I.: Scaling-up social learning in small groups with robot supported collaborative learning (RSCL): Effects of learners' prior experience in the case study of planning poker with the robot NAO. Appl. Sci. **13**(7), 4106 (2023). https://doi.org/10.3390/app13074106
8. Klęczek, K., Rice, A., Alimardani, M.: Robots as mental health coaches: a study of emotional responses to technology-assisted stress management tasks using physiological signals. Sensors **24**(13), 4032 (2024). https://doi.org/10.3390/s24134032
9. Edwards, A., Edwards, C., Abendschein, B., Espinosa, J., Scherger, J., Vander Meer, P.: Using robot animal companions in the academic library to mitigate student stress. Libr. HiTech **40**(4), 878–893 (2022). https://doi.org/10.1108/LHT-07-2020-0148
10. Donnermann, M., Schaper, P., Lugrin, B.: Social robots in applied settings: a long-term study on adaptive robotic tutors in higher education. Front. Robot. AI **9**, 831633 (2022). https://doi.org/10.3389/frobt.2022.831633
11. Qu, J., Jarosz, M., Sniezynski, B.: Robot control platform for multimodal interactions with humans based on ChatGPT. Appl. Sci. **14**(17), 8011 (2024). https://doi.org/10.3390/app14178011
12. Baimukashev, D., Kappassov, Z., Varol, H.A.: Shear, torsion and pressure tactile sensor via plastic optofiber guided imaging. IEEE Robot. Autom. Lett. **5**(2), 2618–2625 (2020). https://doi.org/10.1109/LRA.2020.2972876
13. Lin, G.-Y., Jhang, C.-C., Wang, Y.-S.: Factors affecting parental intention to use AI-based social robots for children's ESL learning. Educ. Inf. Technol. **29**(5), 6059–6086 (2024). https://doi.org/10.1007/s10639-023-12023-w
14. Gongor, F., Tutsoy, O.: Doctor robots: design and implementation of a heart rate estimation algorithm. Int. J. Soc. Robot. **14**(6), 1435–1461 (2022). https://doi.org/10.1007/s12369-022-00888-9
15. Chraibi, K., Chaker, I., Dhassi, Y., Zahi, A.: MSAPersonality: a modern standard Arabic dataset for personality recognition. Int. J. Electr. Comput. Eng. IJECE **14**(4), 4498 (2024). https://doi.org/10.11591/ijece.v14i4.pp4498-4507
16. Song, H., Tsiakas, K., Ham, J., Markopoulos, P., Barakova, E.I.: 'How would you score yourself?': The effect of self-assessment strategy through robots on children's motivation and performance in piano practice. Int. J. Soc. Robot. **16**(2), 327–340 (2024). https://doi.org/10.1007/s12369-023-01080-3

17. Yi, E., Park, D.-H.: Potential user segmentation based on expectations of social robots using Q-methodology. IEEE Access **12**, 100295–100315 (2024). https://doi.org/10.1109/ACCESS.2024.3430864
18. Darmawansah, D., Hwang, G.-J.: Effects of robot-based multiple low-stakes assessments on students' oral presentation performance, collective efficacy, and learning attitude. Educ. Technol. Res. Dev. **72**(4), 2013–2039 (2024). https://doi.org/10.1007/s11423-024-10360-2
19. Stapels, J.G., Penner, A., Diekmann, N., Eyssel, F.: Never trust anything that can think for itself, if you can't control its privacy settings: the influence of a robot's privacy settings on users' attitudes and willingness to self-disclose. Int. J. Soc. Robot. **15**(9–10), 1487–1505 (2023). https://doi.org/10.1007/s12369-023-01043-8
20. Pinney, J., Carroll, F., Newbury, P.: Human-robot interaction: the impact of robotic aesthetics on anticipated human trust. PeerJ Comput. Sci. **8**, e837 (2022). https://doi.org/10.7717/peerj-cs.837
21. Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., Tanaka, F.: Social robots for education: a review. Sci. Robot. **3**(21), eaat5954 (2018). https://doi.org/10.1126/scirobotics.aat5954
22. Maggioni, M.A., Rossignoli, D.: If it looks like a human and speaks like a human … communication and cooperation in strategic human-robot interactions. J. Behav. Exp. Econ. **104**, 102011 (2023). https://doi.org/10.1016/j.socec.2023.102011
23. Khairy, D., Alkhalaf, S., Areed, M.F., Amasha, M.A., Abougalala, R.A.: An algorithm for providing adaptive behavior to humanoid robot in oral assessment. Int. J. Adv. Comput. Sci. Appl. **13**(9) (2022). https://doi.org/10.14569/IJACSA.2022.01309119
24. Okawa, K., Jimenez, F., Akizuki, S., Yoshikawa, T.: Editorial Office: Proposal of learning support model for teacher-type robot supporting learning according to learner's perplexed facial expressions. J. Robot. Mechatron. **36**(1), 168–180 (2024). https://doi.org/10.20965/jrm.2024.p0168
25. Donnermann, M., et al.: Social robots and gamification for technology supported learning: an empirical study on engagement and motivation. Comput. Hum. Behav. **121**, 106792 (2021). https://doi.org/10.1016/j.chb.2021.106792
26. Lopez-Caudana, E., López-Orozco, C.F., Mendoza Bárbara, C., Baltazar Reyes, G.E., Ponce, P., Chong-Quero, J.E.: Physical therapy using robotics: a project-based learning experience for undergraduate students. Australas. J. Educ. Technol. **37**(5), 32–42 (2021). https://doi.org/10.14742/ajet.7139
27. Page, M.J., et al.: Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas. Rev. Esp. Cardiol. **74**(9), 790–799 (2021). https://doi.org/10.1016/j.recesp.2021.06.016
28. Rasouli, S., Ghafurian, M., Nilsen, E.S., Dautenhahn, K.: University students' opinions on using intelligent agents to cope with stress and anxiety in social situations. Comput. Hum. Behav. **153**, 108072 (2024). https://doi.org/10.1016/j.chb.2023.108072
29. Mohan, M., Nunez, C.M., Kuchenbecker, K.J.: Closing the loop in minimally supervised human–robot interaction: formative and summative feedback. Sci. Rep. **14**(1), 10564 (2024). https://doi.org/10.1038/s41598-024-60905-x
30. Johal, W., Gatos, D., Yantac, A.E., Obaid, M.: Envisioning social drones in education. Front. Robot. AI **9**, 666736 (2022). https://doi.org/10.3389/frobt.2022.666736
31. Istenic, A., Bratko, I., Rosanda, V.: Are pre-service teachers disinclined to utilise embodied humanoid social robots in the classroom?. Br. J. Educ. Technol. **52**(6), 2340–2358 (2021). https://doi.org/10.1111/bjet.13144
32. Jakonen, T., Jauni, H.: Managing activity transitions in robot-mediated hybrid language classrooms. Comput. Assist. Lang. Learn. **37**(4), 872–895 (2024). https://doi.org/10.1080/09588221.2022.2059518

33. Csaba, H., Adam, C.: Analysis of human-robot interactions as a sustainability factor. Chem. Eng. Trans. **107**, 85–90 (2023). https://doi.org/10.3303/CET23107015
34. Björling, E.A., Ling, H., Bhatia, S., Matarrese, J.: Sharing stressors with a social robot prototype: what embodiment do adolescents prefer? Int. J. Child-Comput. Interact. **28**, 100252 (2021). https://doi.org/10.1016/j.ijcci.2021.100252
35. Baksh, F., Zorec, M.B., Kruusamäe, K.: Open-source robotic study companion with multimodal human-robot interaction to improve the learning experience of university students. Appl. Sci. **14**(13), 5644 (2024). https://doi.org/10.3390/app14135644
36. Papadopoulos, I., Lazzarino, R., Miah, S., Weaver, T., Thomas, B., Koulouglioti, C.: A systematic review of the literature regarding socially assistive robots in pre-tertiary education. Comput. Educ. **155**, 103924 (2020). https://doi.org/10.1016/j.compedu.2020.103924
37. Lopez-Caudana, E.O., George-Reyes, C.E., Avello-Martínez, R.: Developing the skills for complex thinking research: a case study using social robotics to produce scientific papers. Front. Educ. **9**, 1322727 (2024). https://doi.org/10.3389/feduc.2024.1322727
38. Fengmei, H.: A Q-learning approach for optimizing the impact of musical education using virtual reality and social robots. Mob. Netw. Appl. (2024). https://doi.org/10.1007/s11036-024-02375-z
39. Ahmad, M.I., Refik, R.: 'No Chit Chat!' a warning from a physical versus virtual robot invigilator: which matters most? Front. Robot. AI **9**, 908013 (2022). https://doi.org/10.3389/frobt.2022.908013
40. Elmimouni, H., Rode, J.A., Šabanović, S.: Articulation work for supporting the values of students attending class via telepresence robots. Int. J. Hum.-Comput. Stud. **190**, 103318 (2024). https://doi.org/10.1016/j.ijhcs.2024.103318
41. Algerafi, M.A.M., Zhou, Y., Alfadda, H., Wijaya, T.T.: Understanding the factors influencing higher education students' intention to adopt artificial intelligence-based robots. IEEE Access **11**, 99752–99764 (2023). https://doi.org/10.1109/ACCESS.2023.3314499
42. Reich-Stiebert, N., Eyssel, F., Hohnemann, C.: Exploring University students' preferences for educational robot design by means of a user-centered design approach. Int. J. Soc. Robot. **12**(1), 227–237 (2020). https://doi.org/10.1007/s12369-019-00554-7
43. Al, G.A., Hedworth, N., Tilley, D., Ahmed, S., Afeawo, R., Martinez-Hernandez, U.: BEATRIX: an open source humanoid head platform for robotics teaching and research. HardwareX **20**, e00591 (2024). https://doi.org/10.1016/j.ohx.2024.e00591
44. Bertacchini, F., Scuro, C., Pantano, P., Bilotta, E.: A project based learning approach for improving students' computational thinking skills. Front. Robot. AI **9**, 720448 (2022). https://doi.org/10.3389/frobt.2022.720448
45. De Rooij, A., Broek, S.V.D., Bouw, M., De Wit, J.: Co-creating with a robot facilitator: robot expressions cause mood contagion enhancing collaboration, satisfaction, and performance. Int. J. Soc. Robot. **16**, 2133–2152 (2024). https://doi.org/10.1007/s12369-024-01177-3
46. Velentza, A.-M., Fachantidis, N., Lefkos, I.: Human-robot interaction methodology: robot teaching activity. MethodsX **9**, 101866 (2022). https://doi.org/10.1016/j.mex.2022.101866
47. Stapels, J.G., Eyssel, F.: Robocalypse? Yes, Please! the role of robot autonomy in the development of ambivalent attitudes towards robots. Int. J. Soc. Robot. **14**(3), 683–697 (2022). https://doi.org/10.1007/s12369-021-00817-2
48. Shimaya, J., Yoshikawa, Y., Ogawa, K., Ishiguro, H.: Robotic question support system to reduce hesitation for FACE- TO- FACE questions in lectures. J. Comput. Assist. Learn. **37**(3), 621–631 (2021). https://doi.org/10.1111/jcal.12511
49. Virkus, S., et al.: The potential of telepresence in libraries: students' perspectives. Libri **74**(2), 105–118 (2024). https://doi.org/10.1515/libri-2023-0082

Artificial Intelligence in Higher Education: Student Perceptions of the Adoption and Integration in Ghana, West Africa

Eric Owusu[1(✉)], Ralitsa Debrah[2], and Helina Oladapo[3]

[1] Department of Computing Sciences, State University of New York, Brockport, 350 New Campus Drive, Brockport, NY 14420, USA
eowusu@brockport.edu

[2] Department of Communication Design, Kwame Nkrumah University of Science and Technology, KNUST, CABE, FOA, DECODE, Kumasi, Ghana
rddebrah@knust.edu.gh

[3] Department of Computer and Information Technology, Miami University, 1601 University Blvd., Hamilton, OH 45011, USA
oladaph@miamioh.edu

Abstract. This study investigates the integration of artificial intelligence (AI) in higher education institutions (HEIs) within emerging contexts, using a technology-focused university in Ghana as a case study. In an era of rapid AI advancements, HEIs face unique challenges and opportunities in adopting these technologies to enhance educational outcomes. The research focuses on students' perceptions, preparedness, and support for AI implementation, revealing a generally positive outlook toward increased AI adoption in academic settings. A mixed-methods case study approach was employed, combining survey responses from 200 participants to explore key determinants influencing AI adoption. The study addresses three core questions: (i) To what extent is AI currently integrated into higher education curricula in Ghana? (ii) What factors facilitate or hinder AI adoption in these institutions? (iii) How does AI impact teaching strategies, student engagement, and learning outcomes? Findings indicate that while students are receptive to AI integration, successful implementation requires careful consideration of cultural sensitivity, accessibility, inclusivity, ethical implications, and sustainability. The study provides insights into strategies for equitable AI adoption, ensuring that the benefits of emerging technologies are accessible across Ghanaian HEIs, ultimately contributing to more innovative, inclusive, and effective educational practices.

Keywords: Artificial Intelligence · Adoption · Accessibility · Cultural sensitivity · Equity · Higher education · Inclusive · Integration · Perception

1 Introduction

The implementation of artificial intelligence (AI) in higher education institutions (HEIs) has remained relatively limited compared to other sectors [1]. A notable subset of AI, known as *Generative Artificial Intelligence* (GAI), is designed to produce original content—such as text, audio, or video—by learning from vast datasets [2]. The integration

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 257–271, 2025.
https://doi.org/10.1007/978-3-031-93539-8_17

of AI in educational settings holds significant potential to improve academic writing, personalize learning experiences, streamline research processes, enhance accessibility, and optimize administrative tasks like grading, scheduling, and enrollment [2]. AI in Education (AIED) refers specifically to the application of AI technologies to support and enhance teaching, learning, and decision-making within educational environments [3]. However, several challenges impede the successful implementation of AI in HEIs, particularly in West Africa, where educators often lack the necessary technical expertise [4], coupled with insufficient technological infrastructure [5].

In a study by [6], AIED has been an emerging field in educational technology for approximately three decades. However, many educators continue to struggle with maximizing its potential in teaching and learning environments. In many developing regions, the value of AI in education remains underappreciated, with public policy frameworks still in their infancy. Nevertheless, significant growth in AI adoption is anticipated within the next decade [7]. African nations, in particular, have been slower to adopt AI compared to more developed countries, with limited implementation across educational sectors [8, 9]. Despite this, AI is projected to contribute up to $15.7 trillion to the global economy by 2030 [10].

Over the past 15 years, Africa has witnessed remarkable growth in mobile phone accessibility, which has significantly contributed to economic development [11]. This increase in mobile technology adoption has attracted interest from tech investors eager to explore AI applications across the continent. Countries such as South Africa, Nigeria, Ethiopia, Kenya, and Ghana have made notable strides in integrating AI into their educational systems [12]. However, many other African nations lag, unable to fully harness the benefits of AI due to persistent challenges related to infrastructure, governance, educational systems, and data ecosystems. Overcoming these barriers is essential to unlocking AI's full potential in enhancing educational outcomes across the continent.

2 Research Background

Artificial Intelligence (AI) refers to machine systems designed to replicate human cognitive abilities and perform tasks traditionally requiring human intelligence [9]. The integration of AI in higher education across Africa is influenced by a complex interplay of factors, including infrastructure limitations, cultural attitudes, and the readiness of both educators and institutions. A comprehensive analysis of these factors provides insights into the challenges and opportunities for AI adoption in the region.

2.1 Infrastructure Limitations

Infrastructure remains a significant challenge for the effective adoption and utilization of AI technologies in Africa. Many countries across the continent continue to grapple with limited access to reliable electricity and internet connectivity, which are fundamental for AI systems to function efficiently [8]. Inadequate telecommunications infrastructure, frequent mobile network disruptions, and insufficient power supply hinder efforts to implement AI-driven educational solutions [13]. Without robust infrastructure, the potential for AI to enhance educational outcomes remains constrained.

2.2 Skills and Technology Knowledge

The availability of technical skills is critical for the successful deployment and utilization of AI technologies. However, a pronounced skills gap persists in many African countries. [14] emphasize that the acquisition of technical competencies presents a considerable hurdle to adopting new technologies. Similarly, Bianco identifies the lack of AI-related expertise as a significant barrier, particularly in emerging economies [15]. [16] further underscores the scarcity of skilled professionals capable of effectively implementing and maintaining AI systems, limiting the potential for transformative educational advancements.

2.3 Cultural Considerations

Cultural values, practices, and societal norms play a pivotal role in determining the success or failure of technological innovations. Research conducted in Nigeria and South Africa indicates that cultural factors significantly influence the acceptance and utilization of new technologies [17, 18]. [19] caution that technological solutions designed without considering local cultural dynamics often fail to gain traction in developing regions. As such, it is essential for technology developers and policymakers to integrate culturally sensitive approaches when designing and implementing AI solutions in African HEIs.

2.4 Knowledge Deficiency and Literacy Gaps

AI technology is grounded in complex mathematical and computational principles, requiring a solid foundation in these disciplines for effective implementation. However, low literacy rates and limited digital literacy in several African regions pose significant challenges to AI adoption [20, 21]. [22] highlight the widespread lack of awareness and understanding of AI applications, which further impedes integration efforts. Establishing effective knowledge management systems and investing in educational initiatives to improve AI literacy are essential steps to bridge this gap.

The successful integration of AI into higher education in Africa requires addressing critical infrastructural deficits, closing the technical skills gap, incorporating cultural considerations, and enhancing AI-related knowledge through targeted educational programs. Overcoming these challenges can position African HEIs to fully harness AI's potential in transforming educational delivery, improving student outcomes, and contributing to broader socio-economic development.

3 Methodology

This research employed a mixed-methods case study methodology survey to investigate the reasons for AI integration, student involvement, encountered challenges, and effective implementation techniques. Data was gathered and analyzed with responses from 200 participants in Ghana, concentrating on the following study questions: (i) To what extent is AI currently integrated into higher education curricula in Ghana? (ii) What factors facilitate or hinder AI adoption in these institutions? (iii) How does AI impact teaching strategies, student engagement, and learning outcomes?

The respondents in this study were primarily drawn from creative disciplines, as they were readily available and demonstrated frequent use of technology in their academic and practical work, making the study's questions particularly relevant to their experiences. To align with the research objective of understanding AI adoption across diverse academic fields, we employed a combination of purposive and random sampling techniques. This approach enabled the inclusion of students from non-creative disciplines, providing broader insights into the perceptions and readiness for AI integration in teaching and learning within HEIs. The integration of artificial intelligence in educational environments encounters many critical problems, namely, readiness and adoption, and the inadequacy of appropriate technology infrastructure.

4 Results and Discussions

The majority of respondents were youths aged 18 to 25 years (75.0%), with 48.5% falling within the 18 to 21-year age range. Males made up the majority of the sample (73.5%), and 58.0% of the respondents were in their second year of study. Additionally, most respondents (88.0%) were full-time students (Table 1). Communication design was the most common academic major, representing 52.0% of respondents, followed by visual communication at 16.0% (Fig. 1).

65.0% of respondents had previously used AI-based tools in their learning, and 89.0% were estimated to be current users of the tool at the time of data collection. The three top commonly used AI types were ChatGPT (50.0%), Quillbot (14.6%) and Photoshop (10.1%). Just over half of the respondents (51.5%) described their current level of experience with AI technologies as intermediate (see Figs. 2, 3, 4 and 5).

Perceived Challenges or Concerns with Integrating AI into Higher Education. The respondents identified several key challenges or concerns regarding the integration of AI into higher education. These included limited internet connectivity ($M = 26.5 \pm 1.5$), the cost of implementation ($M = 6.3 \pm 1.6$), government policies and support ($M = 26.3 \pm 1.5$), the availability of data centers and cloud services ($M = 26.1 \pm 1.7$), and data protection and privacy laws ($M = 26.0 \pm 1.6$) (Table 2).

The results from Table 3 revealed that the majority (60.5%) of respondents believed that AI technologies are a supporting factor that enhances students' educational outcomes and 51.0% of them were somewhat concerned about ethical issues related to AI usage in higher education.

The majority of respondents believe that AI technologies enhance students' academic performance, with 60.5% indicating that these tools support learning through personalized instruction, adaptive feedback, intelligent tutoring systems, and data-driven insights. This positive outlook suggests a readiness for educational institutions to adopt AI-driven tools while highlighting the need to explore specific benefits, such as improved learning efficiency and increased engagement. Ethical concerns about AI usage in higher education are a key insight, with 51.0% of respondents expressing concerns about bias, data privacy, and transparency. These concerns stem from unequal treatment of student groups, handling sensitive information, and transparency in AI algorithms' decision-making. Institutions must address these ethical concerns proactively, implementing transparent AI systems, educating stakeholders, and establishing clear data privacy protocols.

Table 1. Socio-demographic characteristics of respondents (N = 200).

Variable	Frequency (N)	Percentage (%)
18–21	97	48.5
22–25	53	26.5
26–29	11	5.5
30–33	14	7.0
34–37	9	4.5
38–41	4	2.0
42–45	9	4.5
50–53	2	1.0
58–61	1	0.5
Gender		
Male	147	73.5
Female	53	26.5
Class year		
First Year	2	1.0
Second Year	116	58.0
Third Year	7	3.5
Fourth Year	27	13.5
Graduate Student	48	24.0
Enrollment Status		
Part-Time	24	12.0
Full-Time	176	88.0

Readiness and Support for AI Adoption. Table 4 below shows responses from the participants regarding their readiness and support for AI adoption. More than half (55.5%) of the respondents somewhat believe that their institution is ready to adopt AI technologies. Those who assumed that necessary resources (e.g., funding, training, technical support) were somewhat available for AI integration were 57.5%, and 61.5% accepted that training and professional development in AI technologies for faculty and staff was very important. In the same vein, 64.5% of respondents believe that students would react positively to the increased use of AI in their learning experiences.

Implementation and Future Perspectives (IFP). Respondents' most common rated areas to ensure implementation of AI within learning institutions are research (33.5%), student support services (21.2%), and administrative tasks (15.2%) followed by curriculum development (14.3%) in (Table 5). Those who foresee long-term impacts of AI integration on higher education in the area of research acceleration make up 38.6%, followed by 34.1% in personalized learning (Table 6).

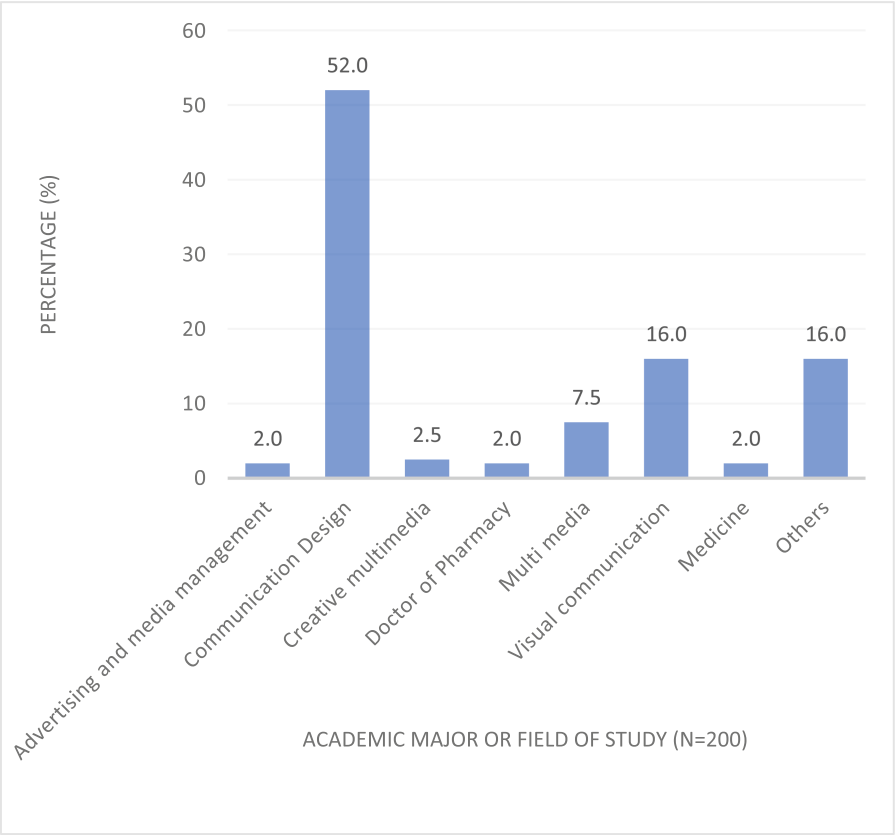


Fig. 1. Respondents' academic major or field of study.

The data highlights the critical role of AI in advancing research productivity and improving personalized learning experiences. To maximize these benefits, educational institutions should adopt a strategic, evidence-based approach to integrate AI into curriculum development and research practices. Investments in AI tools, faculty training, and ethical frameworks will ensure responsible, effective implementation that aligns with long-term educational goals.

Among respondents' suggestions for effective integration of AI technologies into higher education were: It should be adopted in all learning institutions to enrich learning and research experiences (30.6%); There should be capacity building training for staff and students on its adoption and utilization (16.1%); It should be able to improve students' learning capacity and skills (11.3%) and It should be accessible, fast and time saving to all users (9.7%) (Table 9).

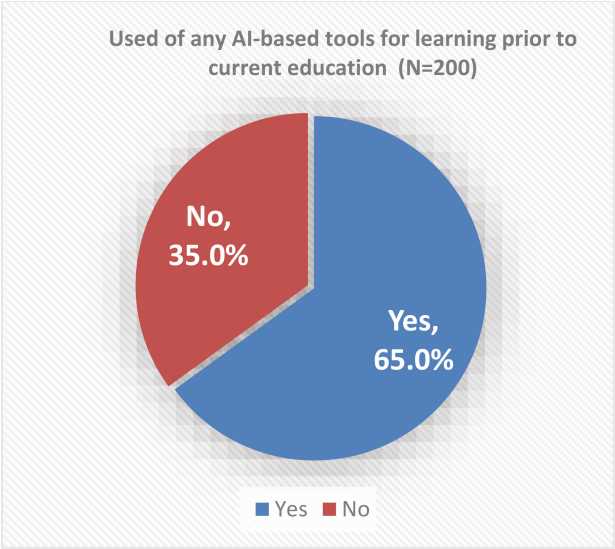


Fig. 2. Respondents' previous use of AI technology.

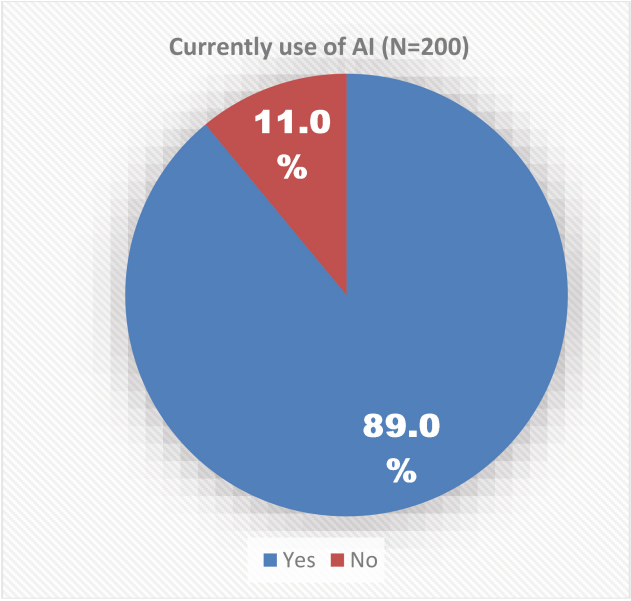


Fig. 3. Respondents currently using AI technology.

The majority (77.0%) of respondents reported their willingness to adopt and support the integration of AI technologies in their studies (Table 8).

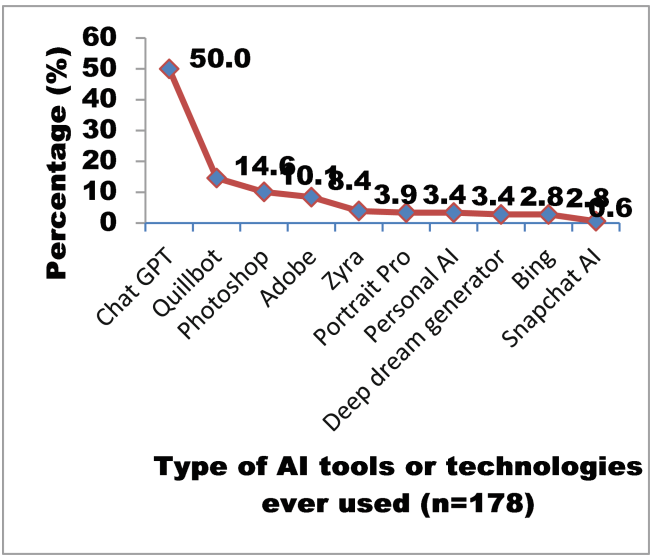


Fig. 4. Type of AI tools or technologies ever used by respondents.

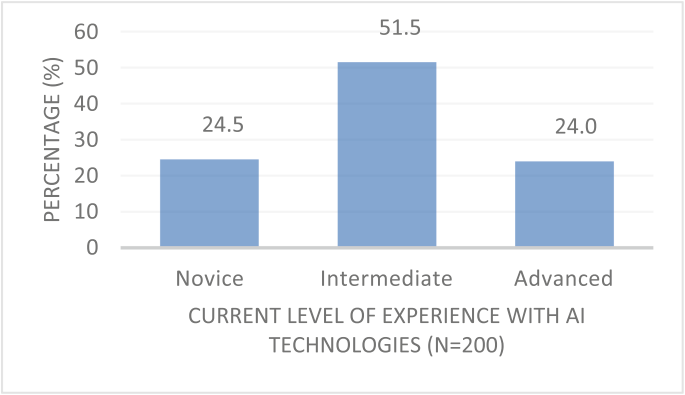


Fig. 5. Experience with AI

Logistic Regression Analysis for Factors Influencing the Use of AI. To identify the principal determinants or predictors of the utilization of AI tools in higher learning institutions, a logistic regression model in (Table 9) was developed using variables constituted challenges to AI-based tools, as determined by the bivariate analysis. The Result showed that variables that retained their significant association with respondents' current use of AI were: Limited internet connectivity (OR = −0.199; P = 0.049); Accessibility (oR = 0.280; P = 0.001) and Inclusion and Equity (OR = −0.200; P = 0.026). These were the predictors of AI usage. However, accessibility was 3 times more likely a factor to influence utilization of AI than other predictors in this category inclusion and equity were half as likely to be influencing factors in the use of AI.

Table 2. Perceived challenges or concerns with integrating AI into higher education (N = 200).

Perceived as the biggest challenges or concerns with integrating AI into higher education	Not Challenging at all Freq.(%)	Very Challenging Freq.(%)	Extremely Challenging Freq.(%)	Mean(±)
Limited Internet connectivity	53(26.5)	99(49.5)	48(24.0)	26.5 ± 1.5
Accessibility	117(58.5)	70(35.0)	13(6.5)	25.3 ± 1.6
Data centers, cloud services	76(38.0)	86(43.0)	38(19.0)	26.1 ± 1.7
Cost of implementation	65(32.5)	89(44.5)	46(23.0)	26.3 ± 1.6
Skills and expertise	118(59.0)	75(37.5)	7(3.5)	25.3 ± 1.5
Government Policies and Support	60(30.0)	110(55.0)	30(15.0)	26.3 ± 1.5
Data Protection and Privacy Laws	75(37.5)	102(51.0)	23(11.5)	26.0 ± 1.6
Education and Training	121(60.5)	56(28.0)	23(11.5)	25.3 ± 1.6
Digital Literacy	107(53.5)	66(33.0)	27(13.5)	25.5 ± 1.7
Cultural Sensitivity	89(44.5)	98(49.0)	13(6.5)	25.7 ± 1.6
Inclusion and Equity	102(51.0)	90(45.0)	8(4.0)	25.5 ± 1.6

(±) – *Standard deviation.*
Composite mean: 25.78 ± 1.58
Symbols: - = < 25.78; + = ≥ 25.78

The respondents identified several benefits of AI technology, including its ease of use, speed and efficient (32.3%). Other perceived advantages mentioned were its role as a relatively advanced tool for academic growth (10.5%), its ability to broaden knowledge and improve user proficiency (8.3%), its capacity to transform ideas and thoughts into reality (7.9%, its potential to analyze student data and provide tailored learning recommendations (7.4%), its usefulness in facilitating learning (5.7%), and its time-saving nature, along with its ability to simply explanation for better understanding (3.5%) (Tables 7 and 10).

Table 3. Effectiveness and ethical concern for AI technology (N = 200).

Statement	Frequency	Percentage (%)
Supporting factors – How effective do you believe AI technologies are in enhancing your educational outcomes?		
Not Effective at all	22	11.0
Somewhat Not Effective	57	28.5
Extremely Effective	121	60.5
Ethical Concern – How concerned are you about ethical issues related to AI use in higher education?		
Not Concerned	20	10.0
Somewhat Concerned	102	51.0
Very Concerned	78	39.0

Table 4. Readiness and Support for AI Adoption (RSAI) (N = 200).

Statement	Frequency	Percentage (%)
How ready do you believe your institution is to adopt AI technologies?		
Not ready	18	9.0
Somewhat ready	111	55.5
Very ready	71	35.5
To what extent are the necessary resources (e.g., funding, training, technical support) available for AI integration?		
Not available	57	28.5
Somewhat available	115	57.5
Fully available	28	14.0
How important is training and professional development in AI technologies for faculty and staff?		
Not important	13	6.5
Somewhat important	64	32.0
Very important	123	61.5
How do you think students would react to the increased use of AI in their learning experiences?		
Negative	8	4.0
Neutral	63	31.5
Positive	129	64.5
Total	200	100.0

Table 5. Implementation areas for AI within learning institution(N = 200)

Variable	Frequency	Percentage (%)
Supposed most potential area for AI to be implemented within institution as by respondents		
Research	185	33.5
Team	63	11.4
Curriculum development	79	14.3
Student support services	117	21.2
Administrative tasks	84	15.2
If not listed, please add area 1	15	2.7
If not listed, please add area 2	6	1.1
If not listed, please add area 3	4	0.7
Total+	553	100.0

+ − *Multiple responses.*

Table 6. Future perspectives of AI in learning institution (N = 200).

Variable	Frequency	Percentage (%)
Long-term impacts foresee the integration of AI having on higher education		
Personalized learning	121	34.1
Improve operational efficiency	97	27.3
Accelerate research	137	38.6
Total+	355	100.0

+ − *Multiple responses.*

Table 7. Suggestions for effectively integrating AI technologies (N = 200).

Suggestions for effectively integrating AI technologies into higher education	Frequency	Percentage (%)
It should be adopted in all learning institutions to enrich learning and research experiences	19	30.6
There should be capacity building training for staff and students on its adoption and utilisation	10	16.1
It should be able to improve students learning capacity and skills	7	11.3
It should be accessible, fast and time saving to all users	6	9.7
There should be provision of resources to cater for its use	4	6.5
Its usage should be monitored and constantly evaluated for improvement	4	6.5

(*continued*)

Table 7. (*continued*)

Suggestions for effectively integrating AI technologies into higher education	Frequency	Percentage (%)
Its usage should be made easier, understood and users friendly	4	6.5
There should be policy towards its adoption and usage	3	4.8
It should be made available for all users	2	3.2
There should be restriction on its usage to students	2	3.2
There should be research tolerance	1	1.6
Total*	62	100.0

**The total is equal to 200 because no suggestion has been deleted*

Table 8. Willingness to adopt and support the integration of AI technologies in studies (N = 200).

Willingness to adopt and support the integration of AI technologies in studies	Frequency	Percentage (%)
Neutral	46	23.0
Willing	154	77.0
Total	355	100.0

Table 9. Logistic regression analysis for factors influencing the use of AI.

Variable	β	S.E	Odds Ratio (OR)	Sig. p = value
(Constant)	0.908	0.098		.000
Limited Internet connectivity	−0.088	0.044	−0.199	.049*
Accessibility	0.142	0.044	0.280	.001*
Data centers, cloud services	−0.012	0.044	−0.029	.777
Cost of implementation	0.031	0.047	0.073	.515
Skills and expertise	0.067	0.041	0.120	.104
Government Policies and Support	0.058	0.040	0.121	.146
Data Protection and Privacy Laws	0.060	0.043	0.125	.158
Education and Training	0.070	0.047	0.154	.138
Digital Literacy	−0.059	0.043	−0.134	.168

(*continued*)

Table 9. (*continued*)

Variable	β	S.E	Odds Ratio (OR)	Sig. p = value
Cultural Sensitivity	−0.015	0.042	−0.029	.725
Inclusion and Equity	−0.109	0.049	−0.200	.026*

S.E. = Standard Error; Ref. = Reference category; * = Significant at $p < 0.05$

Table 10. Benefits of AI within learning institution (N = 200).

Benefit response	Frequency	Percentage (%)
Easy use, fast and efficient	74	32.3
Pretty high technology in academic advancement	24	10.5
It helps in broadening your knowing and makes you more perfect in the use of it	19	8.3
Bringing idea, taught into reality	18	7.9
AI can analyse student data and provide tailored recommendations and resources to enhance their learning experience	17	7.4
Help facilitate learning	13	5.7
It saves time Makes explanations easy to comprehend, since it breaks it down well	8	3.5
It helps students to have easy access to source of information	7	3.1
It narrows a researcher to the exact information needed to execute project	6	2.6
It creates easy and effective communication	6	2.6
Convenient and simple	5	2.2
It will help creative to work faster	5	2.2
It can streamline administrative processes	4	1.7
AI technology enhances designs or projects of the artist	4	1.7
It is good but it makes me a little lazy	4	1.7
AI technologies enhance the quality and overall appearance of a finished project or work	4	1.7
These are technologies that help manipulate images	3	1.3
It speeds the thinking process	3	1.3
The use of artificial intelligence in education can provide many benefits	3	1.3

(*continued*)

Table 10. (*continued*)

Benefit response	Frequency	Percentage (%)
It helps reducing plagiarism	2	0.9
Total+	229	100.0

+− *Multiple responses*

5 Conclusion

The study reveals that AI is transforming higher education in Ghana, with 89% of respondents using AI tools, including ChatGPT, and QuillBot. However, institutions need more infrastructure improvements and faculty training to fully utilize AI in education. Factors such as accessibility, limited internet connectivity, and perceived lack of inclusion, as well as age, class year, and enrollment status, play a role in the adoption of AI tools. Challenges include limited internet connectivity, implementation costs, and lack of government policies and support. Institutional factors like internet access, funding, and demographic characteristics also play critical roles in determining adoption levels. With personalized learning and research facilitation that enhance learning experiences, AI technology is changing how teachers teach, how engaged their students are, and how well their students learn in Ghana's schools. However, infrastructure limitations and concerns around data privacy hinder the full potential of AI in education.

Beyond efficiency, AI also enhances accessibility, making education more inclusive. AI-based platforms can accommodate different learning styles, including visual, auditory, and kinesthetic learners, ensuring that all students feel supported. For those with disabilities, AI technologies like speech-to-text, text-to-speech, and virtual assistants can provide crucial accommodations, improving their ability to fully participate in coursework. Simulations and virtual labs also utilize AI, enabling students to apply their acquired knowledge in a practical environment. This creates dynamic, hands-on experiences that keep students more interested. As a result, students are empowered to take control of their academic journeys with greater autonomy, ultimately contributing to higher motivation, satisfaction, and academic success.

The findings also reveal that AI technologies positively influence teaching approaches by enabling personalized learning, enhancing engagement, and broadening academic capabilities. However, the full potential remains untapped due to infrastructure limitations and concerns around data privacy. In the future, the sample will be expanded even more to include views from a wider range of fields. This will help researchers get a fuller picture of the factors that affect how AI is used in different educational settings.

References

1. Salas-Pilco, S.Z., Xiao, K., Hu, X.: Correction: Salas-Pilco et al. artificial intelligence and learning analytics in teacher education: a systematic review. Educ. Sci. 12, 569. Educ. Sci. **13**(9), 897 (2023). https://doi.org/10.3390/educsci13090897
2. Aderibigbe, A.O., et al.: Artificial intelligence in developing countries: bridging the gap between potential and implementation. Comput. Sci. IT Res. J. **4**(3), 185–199 (2023)

3. Al-Turjman, F., Shankar, P.: Integrating AI and learning analytics for personalized education systems. In: Proceedings of the 25th International Conference on Artificial Intelligence in Education (2024)
4. Chen, Y., Panigrahi, P.: Sustainable AI education: addressing diversity and equity through curriculum innovation. J. Educ. Technol. Soc. (2023)
5. Latif, G., Ben Brahim, G., Abdelhamid, S.E., Alghazo, R., Alhabib, G.: Learning at your fingertips: an innovative IoT-Based AI-powered braille learning system. Appl. Syst. Innov. **6**(5), 91 (2023). https://doi.org/10.3390/asi6050091
6. Akinwalere, S.N., Ivanov, V.: Artificial intelligence in higher education: challenges and opportunities. Border Crossing **12**(1), 1–15 (2022)
7. Pedro, F., et al.: Artificial intelligence in education: challenges and opportunities for sustainable development (2019)
8. Ade-Ibijola, A., Okonkwo, C.: Artificial intelligence in Africa: emerging challenges. In: Eke, D.O., Wakunuma, K., Akintoye, S. (eds.) Responsible AI in Africa. Social and Cultural Studies of Robots and AI, pp. 101–117. Palgrave Macmillan, Cham (2023). https://doi.org/10.1007/978-3-031-08215-3_5
9. Kolog, E.A., et al.: Rethinking the implementation of artificial intelligence for a sustainable education in Africa: challenges and solutions. In: Management and Information Technology in the Digital Era, vol. 29, pp. 27–46. Emerald Publishing Limited (2022)
10. Rao, A.S., Verweij, G.: Sizing the prize: what's the real value of AI for your business and how can you capitalise? (2017). https://preview.thenewsmarket.com/Previews/PWC/DocumentAssets/476830.pdf
11. Aker, J.C., Mbiti, I.M.: Mobile phones and economic development in Africa. J. Econ. Perspect. **24**(3), 207–232 (2010)
12. Gadzala, A.: Coming to life: artificial intelligence in Africa. Atlantic Council–Africa Centre, ISSUU (2018)
13. Oyeyemi, S.O., Wynn, R.: Giving cell phones to pregnant women and improving services may increase primary health facility utilization: a case–control study of a Nigerian project. Reprod. Health **11**, 1–8 (2014)
14. Mtega, W.P., et al.: Using mobile phones for teaching and learning purposes in Higher Learning Institutions: the case of Sokoine University of Agriculture in Tanzania (2012)
15. Bianco, M.: Overcoming the Social Barriers of AI adoption (2021)
16. Ajadi, S.: Can AI help tackle the most pressing challenges in developing countries?. GSM Association (2020). https://www.gsma.com/mobilefordevelopment/region/africa/can-ai-help-tackle-the-most-pressing
17. Bankole, F.O., Bankole, O.O., Brown, I.: Mobile banking adoption in Nigeria. Electron. J. Inf. Syst. Developing Countries **47**(1), 1–23 (2011)
18. Lekhanya, L.M.: Cultural influence on the diffusion and adoption of social media technologies by entrepreneurs in rural South Africa. The International Business Economics Research Journal (2013)
19. Kim, H.-W., Gupta, S.: Investigating customer resistance to change in transaction relationship with an internet vendor. Psychol. Mark. **29**(4), 257–269 (2012). https://doi.org/10.1002/mar.20519
20. Mushemeza, E.D.: Opportunities and challenges of academic staff in higher education in Africa. Int. J. High. Educ. **5**(3), 236–246 (2016)
21. Teferra, D.: The international dimension of higher education in Africa: status, challenges, and prospects. High. Educ. Afr. Int. Dimension **5**, 44–79 (2008)
22. Arakpogun, E.O., et al.: Artificial intelligence in Africa: challenges and opportunities. The Fourth Industrial Revolution: Implementation of Artificial Intelligence for Growing Business Success, pp. 375–388 (2021)

A Management Model for Evaluating Scientific Productivity in Chilean Universities: A Case Analysis

María Angélica S. Piñones[1(✉)], María Paz O. Godoy[2], Geraldine Aroca[2], and Isidora Azocar[2]

[1] Magíster en Planificación y Control de Gestión, Universidad de Valparaíso, Valparaíso, Chile
angelica.pinones@uv.cl

[2] Ingeniería en Información y Control de Gestión, Universidad de Valparaíso, Valparaíso, Chile

Abstract. Scientific productivity in universities is a key indicator for assessing their impact on knowledge, innovation, and the socio-economic development of a country. In Chile, the growing emphasis on research has highlighted the need for management tools that enable an accurate and reliable evaluation of scientific productivity. However, the diversity of internal and external factors affecting research output, along with the complexity of management processes, poses a challenge in developing appropriate evaluation models.

This study aims to establish the guidelines for a management model to evaluate scientific productivity in Chilean universities, integrating both their performance in international rankings and the experiences of researchers within their academic environments.

The analysis is structured around two main axes. First, it examines the factors influencing scientific productivity in Chilean universities. Second, it gathers data to develop customer journey maps that reflect researchers' experiences within their institutions, addressing aspects such as access to resources, institutional support, and administrative challenges.

The findings from both axes provide key insights for outlining a management model to diagnose scientific productivity. These guidelines will contribute to the development of a management framework that assesses both institutional efforts and the indicators that directly influence rankings, fostering a sustainable and high-quality research environment.

Keywords: Scientific Productivity · Management Model · Universities

1 Introduction

Scientific productivity serves as a crucial indicator for assessing the impact of universities on knowledge development, innovation, and socioeconomic progress. In Chile, the growing emphasis on scientific research has underscored the need for robust evaluation models that integrate both international ranking criteria and the lived experiences of researchers within their institutions. However, the complexity of influencing factors—both internal and external—and the variety of metrics employed in global rankings present challenges in designing an effective assessment framework.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 272–289, 2025.
https://doi.org/10.1007/978-3-031-93539-8_18

The Chilean higher education system comprises universities, professional institutes, and technical training centers, governed by the General Education Law (Law No. 20.370/2009). The country currently hosts 58 universities, classified into two main groups: Council of Rectors of Chilean Universities (CRUCH) institutions, which receive direct state funding, and private universities, established after the 1981 reform, which rely on alternative funding mechanisms. These institutions play a pivotal role in scientific research, as the majority of Chilean scientific publications originate from university-affiliated faculty members.

Academic excellence remains a central yet controversial issue in higher education, as its evaluation is closely tied to accreditation processes and curriculum assessments (Salas, 2000) [1]. Scientific productivity, traditionally measured by publication output per research-er (Filippo & Fernández, 2002), is widely recognized as a primary indicator of an institution's research performance (Abramo & D'Angelo, 2015) [2]. Beyond publication counts, scientific productivity also encompasses the broader academic activities of researchers, including technology transfer projects, doctoral dissertations, and research collaborations (De Castro Reis, 2017) [3]; (Salomón & Rodríguez, 2007) [4]; (Páez, 2010) [5]. Moreover, scientific research is increasingly viewed as a driver of innovation and societal progress, necessitating its accessibility and practical application (Colciencias, 2016) [6].

In Chile, universities serve as primary hubs for scientific advancement, with government-funded initiatives supporting research productivity (Barra, 2018) [7]. Key evaluation metrics include publications in indexed journals, technology transfer outcomes, regional and national research impact, and patents. Additionally, external institutions such as research centers, public agencies, and private enterprises contribute to knowledge generation (Opazo, Soto, & Martín, 2023) [8].

Several factors influence scientific productivity in universities, including institutional re-search traditions, faculty workload distribution, access to research funding, and public re-search policies (Buela-Casal, Quevedo-Blasco, & Guillén-Riquelme, 2015) [9]; (Mouton et al., 1990) [10]. The evaluation of scientific quality remains a complex challenge due to the continuous expansion of academic disciplines and the need for reliable assessment method-ologies (Salimi, 2017) [11]; (Mårtensson, 2016) [12]. Additionally, systematic reviews re-quire specialized expertise, posing challenges in measuring research impact (Gómez-Morales, 2020) [13]. Faculty specialization, engagement with the scientific community, and incentive systems—both monetary and non-monetary—further shape research output (Gelfert, 2014) [14].

Research funding has a direct correlation with scientific output, as demonstrated in previous studies (Rosenbloom, 2015) [15]; (Benavente, 2012) [16]; (Xia et al., 2021). Universities that effectively allocate resources and foster interdisciplinary collaboration often achieve higher research impact. Additionally, institutional research centers, government-backed funding mechanisms (such as FONDECYT and CONICYT), and access to specialized soft-ware and databases significantly enhance research capabilities (BCN, 2021) [17]; (GOB, 2023) [18].

Given these complexities, developing a comprehensive evaluation model for scientific productivity is essential. By incorporating both quantitative indicators—such as

publication and citation counts—and qualitative aspects—including researcher experiences, collaboration networks, and institutional conditions—this study aims to propose an integrated approach that enables universities to optimize research strategies and enhance their contribution to scientific knowledge and societal progress.

2 Experimental Stage

The objective of this stage is to collect information through a survey administered to researchers at Chilean universities. This survey, designed using a Likert scale, aims to analyze internal and external factors that may significantly influence scientific productivity within these institutions, from the perspective of the researchers themselves.

To achieve this, the survey has been structured based on the following aspects (see Fig. 1):

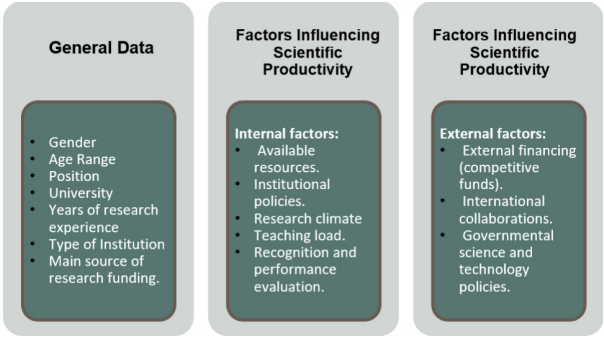


Fig. 1. Survey structure.

The responses obtained through the instrument provide valuable information about the factors influencing scientific productivity in Chilean universities. As part of the sample characterization, Fig. 2 presents the general demographic data of the participating researchers, including the distribution by age, research experience, academic position, and type of university.

Overall, the results based on the sample of researchers from Chilean universities indicate that the participants are adult professionals, mostly over the age of 31. Research experience varies, with a predominance of those with more than 10 years of experience, followed by those with between 5 and 10 years. The majority hold the position of academic/researcher and primarily work in public universities.

2.1 Researcher Experience

This study addresses scientific productivity in Chilean universities, evaluating the internal and external factors that influence the development of research. Through a survey conducted with university researchers, key elements that affect their performance and the conditions under which knowledge production occurs were identified.

Fig. 2. Researcher Characterization Information.

In the survey applied to researchers from Chilean universities, the analysis of internal factors yielded significant results regarding the impact of institutional policies, teaching load, and institutional support on scientific productivity:

The analysis of Fig. 3 reveals that more than 70% of the respondents identified teaching load as one of the main barriers to research development. This limitation is largely due to the lack of available time for scientific activities, which restricts publication output and the development of impactful projects. Furthermore, the high administrative burden associated with research projects represents an additional factor that hinders exclusive dedication to knowledge generation, directly affecting scientific productivity in universities in Chile.

In relation to the availability of infrastructure and funding, Fig. 4 shows that more than 50% (the sum of "Somewhat Disagree" and "Totally Disagree") of respondents perceive that the resources allocated are insufficient for the development of their research. While some researchers have the necessary equipment, the distribution of these resources is not uniform across all institutions. Moreover, the limited frequency with which they receive financial support for their projects constitutes a significant obstacle, hindering the achievement of research objectives and affecting the scientific competitiveness of Chilean universities. These findings highlight the importance of strengthening funding strategies and improving infrastructure conditions to optimize scientific productivity.

In this same context, another internal factor that influences scientific productivity is how universities design and implement their institutional evaluation policies. While universities have systems for evaluating scientific productivity, many researchers perceive that these criteria do not adequately reflect the diversity of activities carried out. There is a noticeable trend of prioritizing quantitative metrics, such as the number of publications, to the detriment of other qualitative indicators, such as community impact or knowledge transfer.

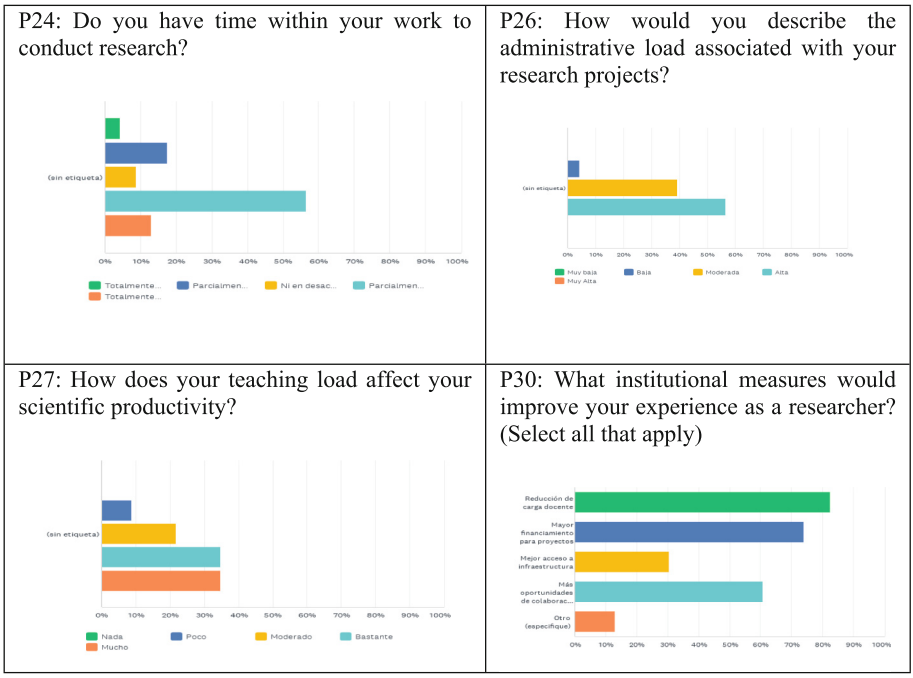


Fig. 3. Survey Results – Internal Factors – Teaching Load.

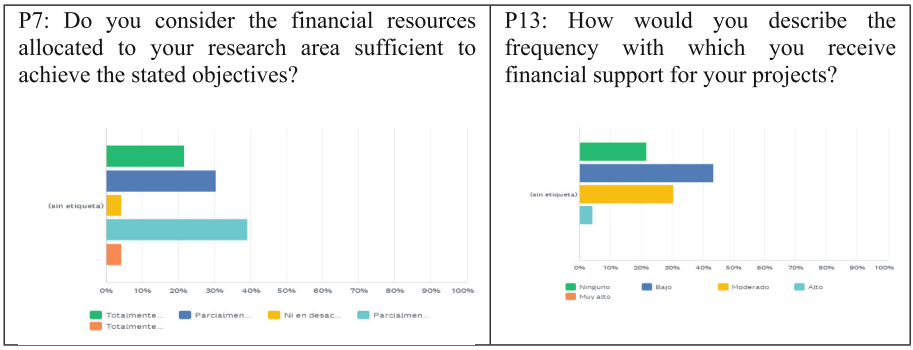


Fig. 4. Survey Results – Internal Factors – Available Resources.

External factors also play a key role in scientific productivity. Among the most relevant findings are (Fig. 5):

Based on the results of question P31 regarding whether the availability of external competitive funds is fundamental for the development of research, it is observed that more than 80% of the participants (the sum of "Somewhat Agree" and "Totally Agree") perceive the availability of external funds as an essential factor to enable, sustain, or enhance their research projects. This finding suggests that the vast majority recognize

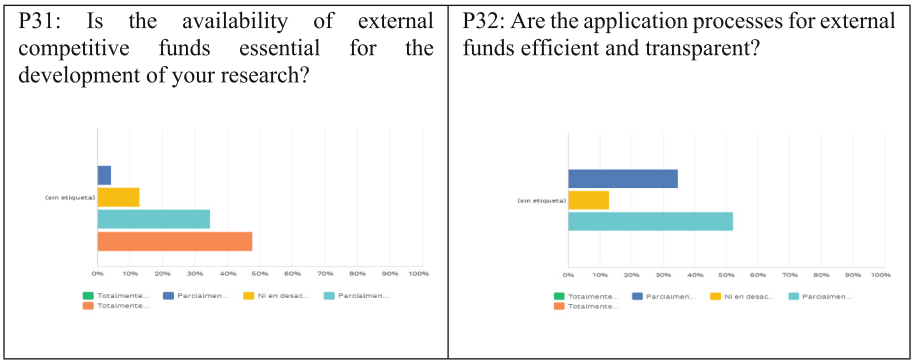


Fig. 5. Survey Results – External Factors – External Funding.

the ability of competitive funds to cover expenses such as equipment, research materials, hiring assistants, and other costs that could otherwise constitute barriers to scientific activity.

The low or non-existent proportion of disagreement and the high average of 4.26 further emphasize the importance of this funding source within the academic community. Therefore, having access to opportunities for external competitive funds appears to be a crucial element in driving scientific productivity and the quality of publications generated in research institutions.

However, it is also mentioned that the application processes are complex and bureaucratic, making it difficult to access these resources.

From the results obtained, it is evident that there is broad agreement on the relevance of external competitive funds for research activities (P31), with a high weighted average (4.26) and more than 80% of respondents expressing partial or total agreement regarding their fundamental role. This suggests that, for the majority, external resources are an essential pillar that facilitates conducting studies, acquiring materials, and disseminating results.

However, the perception of the agility and transparency of the application processes (P32) is less enthusiastic. With an average of 3.17, the evaluation remains in a moderate range: the majority (52.17%) are only "Somewhat Agree," and 34.78% are "Somewhat Disagree." This divided opinion indicates that, although researchers acknowledge the importance of the funds, not all perceive the process of securing them as straightforward or clear.

Together, the findings present a paradox: on one hand, external funds are key to scientific productivity, while on the other, the process of obtaining them may be viewed as limiting due to bureaucracy, lack of feedback, or long delays. Thus, strengthening the transparency and efficiency of these processes emerges as a priority pathway to boost research and optimize the use of these resources.

Regarding international collaboration, Fig. 6 shows that 60.87% of respondents indicate that collaboration with companies is "never" or "infrequent," with only 4.35% considering it "frequent" and another 4.35% "very frequent." The weighted average of 2.30

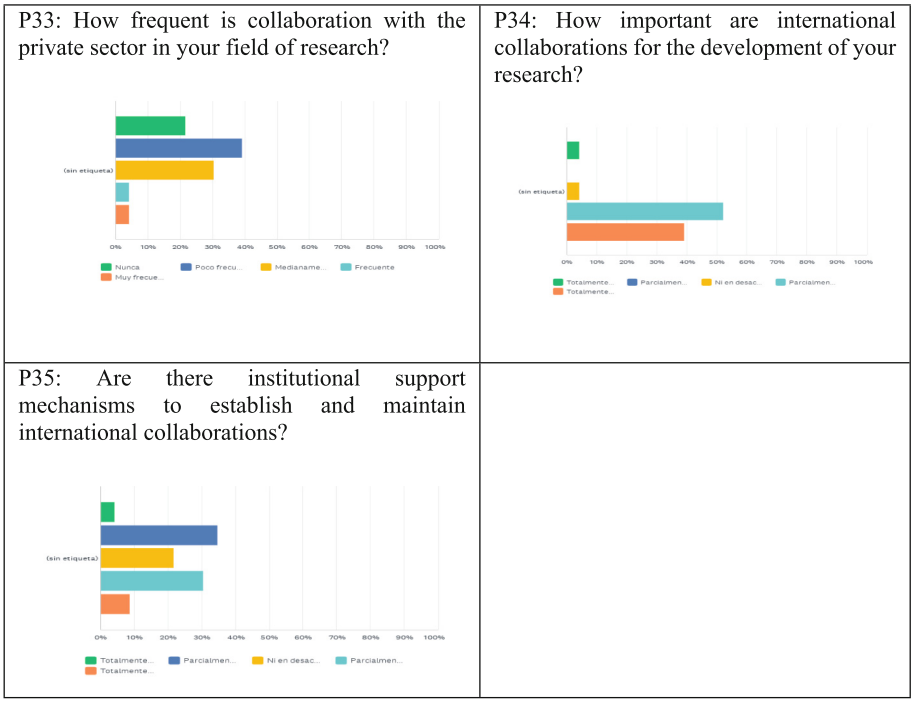


Fig. 6. Survey Results – External Factors – International Collaborations.

is relatively low, indicating that interaction with the private sector is not established as a common practice.

From the results obtained in P34 regarding the importance of international collaborations, 91.3% fall between "partially agree" (52.17%) and "totally agree" (39.13%). The weighted average of 4.22 reflects a highly positive assessment of the relevance of international cooperation. Only 4.35% strongly oppose it, while another 4.35% remain neutral.

These findings highlight a broad consensus on the benefits of engaging with foreign colleagues and research groups, aligning with the global trend of internationalizing research activities to access greater knowledge, networks, and funding. This result contrasts with the low frequency of collaboration with the private sector, demonstrating that the academic community perceives greater value or necessity in expanding internationally rather than engaging with the local business sector.

However, when asked about P35: Institutional support for international collaboration, the weighted average of 3.04 is just above neutrality; 39.13% agree (sum of "partially" and "totally"), 39.13% disagree, and 21.74% remain neutral. This nearly balanced distribution indicates no unanimous opinion: while some perceive institutional support, others believe that conditions or incentives are lacking.

Although most respondents emphasize the importance of international collaborations (P34), not all feel that their institutions provide adequate mechanisms or resources to foster and sustain such partnerships. Stronger internal policies may be required, such as

more agile international relations offices, specific agreements, and funding for mobility or research stays abroad.

Overall, the responses indicate that the research community strongly acknowledges the importance of expanding frontiers, particularly on an international scale, but also recognizes limitations or a lack of institutional impetus for fostering collaboration with both local businesses and global partners. This suggests that while collaboration—especially international—is widely regarded as a means to enhance research quality and reach, strengthening policies and resources to facilitate these alliances is essential to translating collaborative intent into concrete and sustainable projects.

International collaborations are highly valued by researchers, who see them as a key strategy for enhancing scientific productivity. However, institutional challenges limit the establishment and sustainability of these collaborations, such as a lack of logistical support and incentives for internationalization.

Another factor explored among researchers at Chilean universities concerns government policies on science and technology. Some respondents believe that these policies do not create the necessary conditions to strengthen research in Chile. They emphasize the need for greater state investment in science and technology development, as well as the creation of strategies that encourage collaboration between the public and private sectors.

An in-depth analysis of the impact of government support policies on research, such as funding and the creation of specific programs, reveals a significant influence on scientific productivity. Most respondents rank the impact of government policies as "High" or "Very High," while a smaller percentage rates it as "Low" or "Minimal."

An intermediate group may perceive these policies as "Moderate," indicating that while not all researchers feel a decisive impact, they do not dismiss it either. The average rating (on a scale of 1 to 5) generally ranges between 3.5 and 4.0, suggesting that, collectively, participants perceive public policies as a significant factor, albeit not necessarily the most decisive.

This "moderate-high" rating implies that the government's role (e.g., in funding, regulations, research incentives, and ease of importing supplies) can either accelerate or hinder scientific production.

When government policies favor research—through competitive grants, tax exemptions for scientific equipment, support for international mobility, or streamlined permit processes—respondents perceive a direct positive impact on their ability to conduct and publish research. Conversely, bureaucratic barriers or budget cuts result in difficulties in carrying out projects, slowing scientific productivity.

Comparing this dimension with other assessed factors (e.g., external funding, infrastructure, international collaboration), a correlation typically emerges between perceptions of government policies and access to resources. In other words, public policies favorable to R&D tend to facilitate funding acquisition and the establishment of agreements. Likewise, if regulations are complex or lack sufficient incentives, a less conducive environment arises, reflected in lower ratings.

In general, the findings confirm that the academic community considers government policies to exert a significant influence on the ability to conduct quality research and produce scientific publications.

To foster greater scientific productivity, state authorities could reinforce or design policies that encourage research investment, streamline administrative procedures, and promote collaboration programs both nationally and internationally.

This perspective positions government policies as a relevant external factor that can either enhance or constrain research activities. Therefore, monitoring and advocating for regulations and support mechanisms that strengthen the science and innovation ecosystem is crucial.

Additionally, Question 36 ("Are government policies on science and technology favorable to the development of research in your field?") provides a clearer view of both the perceived significance of these policies and their practical evaluation, yielding the following results: 26.09% "Partially disagree," 30.43% "Neither agree nor disagree," 4.78% "Partially agree," 8.70% "Totally agree," and 0% "Totally disagree," with a weighted average of 3.26.

With an average rating of 3.26 (on a scale of 1 to 5), the overall perception is slightly positive but not overwhelmingly so. Most respondents do not see government policies as entirely unfavorable, yet there is no strong consensus rating them as "excellent" or "highly favorable."

The previous question indicated that policies (and their implementation) significantly influence research, perceived as a relevant external factor (with "Moderate/High" impact). However, Question 36 further examines the extent to which these policies are actually "favorable." The result of 3.26 suggests that while public policy influence is recognized, not all respondents feel it provides strong support for their research fields.

The absence of responses in "Totally disagree" indicates no radical rejection of current government policies. Nevertheless, nearly 26% express "Partial disagreement," and 30% remain neutral, reflecting that a significant portion believes policies could be improved or are not sufficiently effective. Only 8.70% view them as entirely favorable, underscoring the need to strengthen state support for research (e.g., increased funding, reduced bureaucracy, and more targeted strategies for emerging fields).

If we compare the previous finding that most respondents acknowledged government policies as an important factor in advancing (or hindering) research with this new indicator on how "favorable" they are perceived to be, we observe a gap between recognized importance and satisfaction with their execution.

In other words, government action is perceived as influential, but not everyone believes the current approach is the most conducive to their projects. Nonetheless, extreme dissatisfaction levels are low (0% "Totally disagree"), suggesting room for improvement:

Refining existing policies to better address the needs of each scientific field, optimizing funding pathways and administrative efficiency, Strengthening communication and engagement channels for the research community.

In summary, the data indicate that government policies on science and technology are viewed as moderately favorable. Acknowledged as an influential factor with little outright disagreement, the critical point lies in enhancing their effectiveness so that state support translates into a more evident and efficient stimulus for research development.

Overall, the results suggest that scientific production in Chilean universities is strongly influenced by a combination of internal factors (motivation, availability of

time, direct institutional support) and external factors (funding, government regulations, private sector engagement, and international collaborations). There is broad consensus on the importance of competitive funding and favorable public policies, although limitations are perceived in the efficiency of application processes and in institutional support for collaboration networks. Furthermore, engagement with businesses appears scarce compared to the high valuation of internationalization. This scenario highlights the need to simultaneously strengthen institutional support for research activities and external conditions (financial, regulatory, and networking) to achieve higher productivity levels and a more robust and sustainable scientific development.

3 Customer Journey Map Reflecting Researchers' Experiences Within Their Universities

The Customer Journey Map (CJM) is a strategic tool that allows for the diagramming of a customer's journey, as well as all their interactions with touchpoints within a given company (Matus, 2020) [19]. For this reason, the use of the customer journey map is becoming increasingly popular due to its usefulness in understanding an organization's customer experience (Rosenbaum, 2016) [20], as it helps identify both positive and negative aspects, as well as critical points that require improvement.

Each point present in the diagram reflects the perception at a given phase, and the classifications include positive experience, which reflects customer satisfaction, while negative experience represents dissatisfaction, raising concerns for strategic management. Finally, a neutral experience is the midpoint between the two.

Figure 7 represents the customer journey map that reflects researchers' experiences within their universities in Chile, regarding the factors influencing scientific productivity, identifying each of the stages, touchpoints, and perceptions, classified into three types of experiences: positive, neutral, and negative.

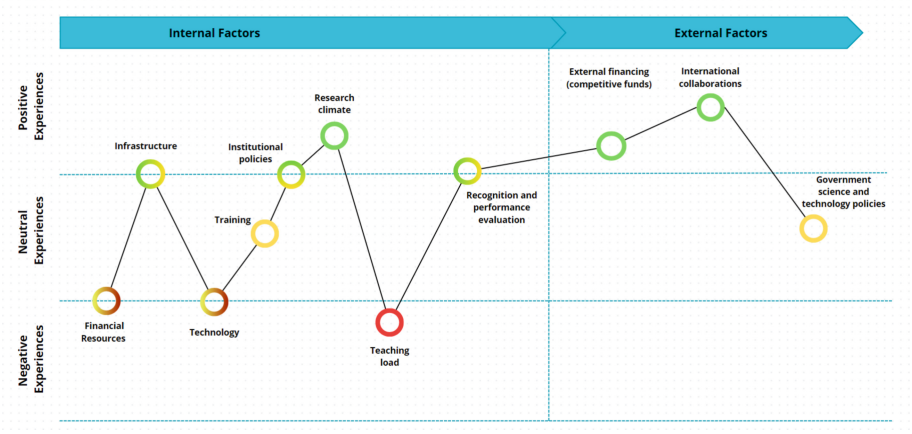


Fig. 7. Customer Journey Map – Researchers in Chilean Universities.

3.1 Description of Touchpoints – Internal Factors

The following is a description of the points of contact identified in Fig. 7 that influence scientific production in Chilean universities.

FPCI 001: Financial Resources. Funding is a crucial component that directly affects the feasibility of initiating research projects and sustaining them through completion. Based on the collected data, many researchers consider the allocation of financial resources insufficient to achieve their study objectives. On one hand, the assessment of whether financial resources are adequate shows an average score of 2.74, bordering on dissatisfaction, with a significant percentage of participants expressing total or partial disagreement. On the other hand, the frequency of financial support is rated even lower (2.17), indicating that most researchers receive little to no regular funding.

Additionally, the survey reveals that nearly two-thirds of respondents identify "Insufficient Funding" as one of the primary factors limiting their productivity. In terms of accessibility, project-specific funding is rated as moderately available (3.17), but other categories, such as funding for attending conferences, are rated even lower (2.52). These limitations hinder the dissemination of research findings and the establishment of collaborations, both of which are crucial for research quality and impact.

Finally, most researchers demand increased funding from their institutions as one of the key measures to improve their experience. In summary, the lack of financial resources not only slows the pace of scientific production but also hinders the consolidation of new research lines, participation in international events, and the adoption of cutting-edge technological infrastructure. The survey results clearly indicate that funding is one of the main constraints on research productivity, as it is perceived as insufficient, infrequent, and only moderately accessible. This directly impacts scientific productivity, project quality, and the ability to disseminate findings nationally and internationally.

FPCI 002: Infrastructure. Infrastructure—including laboratories, offices, technological equipment, libraries, and software—is a fundamental pillar of scientific production, as it largely determines the feasibility and efficiency of conducting research projects. Based on the analyzed data, the overall evaluation of available infrastructure is slightly below neutral satisfaction levels (weighted average of 2.91 out of 5). This suggests that, for many researchers, the current facilities and equipment do not fully meet their needs.

Infrastructure plays a determining role: when it aligns with the requirements of each discipline and is continuously updated, it facilitates research work, increases the chances of publishing in high-impact journals, and promotes connections with external networks. Therefore, the results underscore the urgent need to optimize available spaces and equipment so that researchers can achieve their objectives more effectively and produce higher-quality scientific output.

FPCI 003: Technology. Technology plays an essential role in scientific production by providing tools and platforms that streamline every stage of the research process—from data collection and analysis to results generation and dissemination. When researchers have access to specialized equipment, advanced software, and adequate computing resources, they can optimize experiments, improve analytical accuracy, and reduce execution times. This, in turn, increases both the quantity and quality of publications, as more

robust findings are generated and researchers can respond more swiftly to developments in the scientific community.

However, the collected data reveal that not all researchers have equal access to technological tools. This disparity creates imbalances in productivity: researchers lacking updated technologies face delays and methodological limitations, whereas those with access to cutting-edge technology can progress at a more competitive pace. As a result, achieving solid and sustained scientific production requires institutions to strengthen the provision of equipment and software, promote the adoption of new technologies, and ensure that resources are equitably available across all research groups.

FPCI 004: Training. Training constitutes a fundamental driver for boosting the quality and quantity of scientific production, as it provides researchers with the methodological, management, and dissemination tools necessary to optimize the entire cycle of a project.

The training factor emerges as a key element for the growth of scientific productivity, as it affects both methodological updates and the efficient management of projects and the dissemination of results. Based on the data obtained from the survey, three main findings stand out: The level of institutional support for training (P10) shows a notable percentage of responses indicating "None" (21.74%) and "Low" (39.13%), revealing that almost six out of ten researchers perceive limited or non-existent support in this area. This implies that, although the institution offers some support in other areas, specific training for planning and conducting research (e.g., securing funding, resource management, accountability) is not sufficiently covered.

The second factor is accessibility to professional development opportunities (P19), with a weighted average of 3.35. Academic development opportunities (e.g., courses, diplomas, workshops) are rated as moderately to highly accessible. Approximately half of the respondents describe these opportunities as "Accessible" or "Highly accessible.". However, 4.35% consider them "Not accessible at all" and 17.39% "Barely accessible," suggesting that the training offering does not fully cover the needs of all research fields nor reach the entire academic community equitably.

The third factor is the influence of training on scientific production (P14). In the question regarding the influence of various factors on productivity, training is often rated as an element that has a high or "very high" impact on research performance. This indicates an awareness that continuous training—in methodologies, scientific writing, project management, among others—helps optimize time and resources, improves the quality of results, and strengthens competitiveness at both national and international levels.

In summary, the results reveal a paradox: while training is recognized as a decisive factor in increasing productivity, a greater percentage of researchers perceive that institutional support (particularly in project management) is scarce, and the accessibility of professional development opportunities varies depending on the field or the availability of programs.

FPCI 005: Institutional Policies. Institutional policies play a fundamental role in scientific production within universities, as they establish the normative, strategic, and operational framework that regulates academic research. These policies can influence scientific productivity both positively and negatively, depending on their design, implementation, and alignment with the incentives and needs of the academic body.

In light of these results, several institutional elements directly impact scientific production, such as Clear Research Support Policies (P15), where 86.96% of respondents consider that their university does indeed have clear policies supporting research. This suggests that, at a normative or formal level, the majority perceives the existence of guidelines and directives that could favor research.

Another notable element is the Promotion of Collaboration among Researchers (P16), where the majority does not express strong disagreement, with almost half (43.48%) remaining "neutral"; additionally, only 13.04% are "completely in agreement." This indicates that, while policies exist, not all perceive them as strongly fostering collaboration. The prevailing neutrality may reveal a lack of concrete actions or knowledge of effective collaborative mechanisms.

The existence of clear policies (P15) is a good starting point, but these policies must translate into specific actions that promote collaboration (P16), as collaborative dynamics tend to result in a greater number and quality of publications.

The effectiveness of planning, execution, and monitoring of projects (P17) is essential: if these processes are perceived as merely "moderately effective," research groups are likely to face obstacles in terms of time, resources, or coordination, which can slow down or affect the quality of scientific production.

Having a system for evaluating scientific productivity (P18) can incentivize researchers' performance, provided that the criteria are fair, well-known, and consistent with institutional policies. However, the perception of some lack of clarity in certain cases suggests there may be opportunities to reinforce the dissemination and transparency of such systems.

Together, these factors (institutional policies, promotion of collaboration, efficiency in R&D processes, and clear evaluation systems) impact researchers' motivation, resource management, project development, and ultimately, the quality and quantity of scientific production. Adjusting and strengthening each of these components would contribute to consolidating research outcomes within the institution.

FPCI 006: Research Climate. The work environment within an academic institution plays a decisive role in researchers' ability to generate new knowledge and respond creatively to scientific challenges.

The analysis of the results reveals a partially favorable research climate, where a considerable proportion of participant's report feeling motivated to be creative and innovate in their field and perceive some support from leadership for undertaking research. However, a less positive evaluation is observed regarding collaboration, evidencing an environment where, while creativity is encouraged and some institutional support is offered, it remains essential to strengthen collaborative networks and a culture of teamwork to maximize scientific productivity.

FPCI 007: Academic Workload. The main challenge for university researchers in Chile is balancing teaching, management, and scientific activities. More than 95% of respondents consider that the administrative workload is moderate or high, which hinders scientific productivity. Meanwhile nearly 70% of respondents affirm that teaching workload significantly affects their scientific production, suggesting the need for institutional measures to balance responsibilities.

The results depict a scenario where, although a majority of faculty members acknowledge having some time for research (average 3.57), they also indicate that balancing research with other responsibilities (teaching, management, etc.) is not clearly ensured (average 3.04). This is reflected in the fact that around 39% of respondents disagree (totally or partially) that their institution fosters such a balance, while only 44% agree (totally or partially).

Additionally, the administrative workload is mostly rated as "high" (56.52%) or "moderate" (39.13%), placing the average at 3.52. Finally, another question shows that the vast majority feel that teaching workload "significantly" or "greatly" affects their scientific productivity (69.56% of responses), with an average of 3.96 out of 5.

Collectively, these results suggest that, although there is some available time for research, both excessive teaching responsibilities and administrative procedures continue to pose significant obstacles to achieving optimal scientific productivity. Likewise, the relatively low rating of institutional support for work-life balance indicates that participants do not perceive sufficient institutional support for adequately balancing their research and teaching responsibilities.

FPCI 008: Recognition and Performance Evaluation. Institutional recognition plays a fundamental role in scientific productivity, and its impact can be analyzed from different perspectives, including informal recognition, when academics perceive that their work is valued; monetary and non-monetary incentives; and the creation of supportive work environments, among others. Regarding monetary incentives, the survey results show that no respondent rated them as "not useful at all" or "barely useful"; on the contrary, the majority rated them as "useful" (56.52%) or "very useful" (26.09%), resulting in a weighted average of 4.09. This outcome suggests that monetary incentives are widely seen as an effective driver for encouraging article publication, project participation, and overall scientific productivity.

At the same time, regarding non-financial recognition, 8.70% rated it as "barely useful," and 13.04% as "moderately useful," while the majority (43.48% "useful" and 34.78% "very useful") also valued it positively. This indicates that, beyond monetary compensation, honorary mentions, institutional distinctions, symbolic awards, or additional visibility (for example, through ceremonies or university bulletins) act as relevant motivational factors for academics. They can reinforce self-esteem, reputation, and the desire to maintain or increase scientific productivity.

If the institution adequately recognizes the performance of its researchers with balanced criteria considering the quantity and quality of publications, innovation, social impact, and the training of new scientists, scientific productivity significantly improves.

The survey question regarding whether performance evaluation policies adequately recognize the diversity of research activities yielded an average of 3.09 on a scale of 1 to 5, indicating a moderate level of satisfaction but without strong consensus. In other words, while some perceive a certain level of recognition from the institution, another group believes that not all aspects of research work (such as dissemination, mentoring, collaborative work, or applied research) are valued equitably. This disparity of opinions highlights the need to review and expand evaluation criteria to encompass multiple types of contributions and foster scientific productivity more comprehensively.

Overall, these data suggest that, although a significant proportion of respondents perceive some level of recognition for various forms of research (in line with the findings

on the importance of incentives and non-monetary recognition), there remains a considerable segment that considers the recognition of the diversity of research activities, coverage, and reach of policies insufficient or inadequate. This sense of dissatisfaction demonstrates that, while the implementation of incentives is positively valued, current approaches may not encompass all relevant dimensions of academic work. Thus, it reinforces the idea that improving evaluation and recognition mechanisms—not only through financial incentives but also by promoting inclusive and more detailed criteria regarding different academic contributions—is crucial. This expanded approach would not only contribute to a fairer and more accurate assessment of scientific productivity but also enhance researchers' motivation and satisfaction, thereby boosting scientific productivity in a comprehensive and sustainable manner in the long term.

3.2 Description of Touchpoints – External Factors

FPCE 001: External Funding. External funding plays a key role in the scientific productivity of Chilean universities, as it provides the necessary resources to develop larger-scale and higher-quality research. This type of funding enables the hiring of specialized personnel, the acquisition of advanced equipment, and the execution of fieldwork—elements that often cannot be fully covered by institutional funds alone. Additionally, access to competitive funding fosters the creation of national and international collaboration networks, increasing the visibility and impact of scientific results. The competition for these funds also drives researchers to raise their quality standards and stay up to date in their fields. However, the effectiveness of this funding can be affected by the complexity of application and allocation processes, potentially limiting its ability to maximize scientific productivity.

The data show that the vast majority of researchers highly value the availability of external funding and consider it a decisive element in boosting their scientific output. In fact, the results indicate that approximately 82.6% of respondents (combining "partially agree" and "totally agree") affirm that competitive external funding is essential for carrying out their research, which is also reflected in the weighted average score (4.26 out of 5).

However, when analyzing the survey results regarding whether application processes are agile and transparent, the picture becomes more nuanced. While more than half of the participants (52.17%) "partially" agree that these processes are adequate, 34.78% express some level of disagreement, and no respondent indicates total agreement. This is reflected in a weighted average of 3.17, demonstrating mixed perceptions: although researchers recognize the existence of pathways to access funding, they also identify potential areas for improvement in terms of simplification, speed, clarity of evaluation criteria, and transparency in allocation.

In this context, it is evident that while external funding is fundamental for driving scientific productivity, its effectiveness could be further enhanced through the optimization of application and allocation procedures.

FPCE 002: International Collaboration. International collaborations have a significant influence on the scientific productivity of Chilean universities, as they facilitate access to global knowledge networks, advanced technological resources, and shared

funding. These partnerships allow researchers to participate in large-scale projects, publish in high-impact journals, and increase the visibility of their work at the international level. Additionally, they promote the exchange of ideas and innovative methodologies, strengthening the quality of research. However, inadequate institutional support can limit the ability to fully leverage these opportunities, affecting the potential for scientific growth.

The results indicate that a significant majority of surveyed researchers recognize the importance of international collaborations for the development of their research. A total of 52.17% of participants are partially in agreement, while 39.13% are fully in agreement, summing up to 91.3% who positively value these collaborations. Only 4.35% expressed total disagreement, while another 4.35% remained neutral. The weighted average of 4.22 supports this positive perception, reflecting that international collaborations are seen not only as an opportunity to access broader academic networks but also as a means to improve the quality and impact of research.

However, in contrast to the high valuation of international collaboration, survey results reveal more critical perceptions regarding the support institutions provide for establishing and maintaining these collaborations. Only 8.70% of respondents completely agree that such support exists, while 30.43% partially agree. On the other hand, 34.78% express partial disagreement, and 4.35% are totally in disagreement. The weighted average of 3.04 reflects a mixed perception, close to neutrality, suggesting that while some researchers find institutional support, a significant portion identifies barriers or deficiencies in this area.

FPCE 003: Government Policies. Government policies in science and technology are a fundamental component in shaping the environment in which scientific research develops. In the case of Chilean universities, these policies can influence scientific productivity in various ways, such as through funding and resources, infrastructure and equipment, regulation, and bureaucracy, among others.

Survey results reveal mixed perceptions regarding the influence of government policies on scientific research development. Only 8.70% of respondents fully agree that these policies are favorable, while 34.78% partially agree. However, 26.09% express partial disagreement, and 30.43% remain neutral (neither agreeing nor disagreeing). The weighted average of 3.26 suggests a moderate evaluation, just above neutrality.

This result indicates that while some researchers recognize certain advances or support from public policies, there is a general perception that these are not sufficiently strong or effective in significantly enhancing scientific development. The lack of clear consensus toward a positive assessment reflects potential deficiencies in the formulation, implementation, or scope of current policies.

4 Conclusions and Recommendations

The analysis conducted highlights that scientific productivity in Chilean universities is determined by a combination of internal factors, such as teaching workload, resource availability, and institutional support, as well as external factors, including access to funding, government policies, and international collaboration opportunities. Researchers

emphasize the importance of external funding and international collaborations as key drivers of scientific productivity. However, they also identify significant barriers, such as the complexity of grant application processes and the lack of sufficient institutional support to establish and maintain collaborative networks.

Additionally, government policies in science and technology are perceived as moderately favorable, reflecting the need for greater state commitment in terms of funding, infrastructure, and bureaucracy reduction. The results suggest that, although international rankings provide a useful reference for assessing university competitiveness, their quantitative approach does not fully capture the diversity and real impact of research.

Based on the results obtained, the following recommendations are highlighted:

- It is recommended to simplify the application processes for competitive external funding, reducing bureaucratic burdens and improving feedback for researchers. This would enable more equitable and efficient access to the resources necessary to drive research.
- Universities should develop specific strategies to promote internationalization, including the creation of more agile international relations offices, specific collaboration agreements, and funding programs for research stays and joint projects.
- It is proposed that the government increase investment in science and technology by designing more inclusive policies that address the specific needs of different research areas. Additionally, it is essential to streamline administrative procedures and facilitate access to high-quality infrastructure.
- Universities should review their systems for evaluating scientific productivity by incorporating qualitative metrics that reflect social impact, knowledge transfer, and innovation. This would allow for a fairer and more comprehensive assessment of researchers' performance.
- To optimize scientific productivity, it is crucial to implement measures that enable researchers to balance their teaching responsibilities with research time, reducing administrative workload and promoting policies for exclusive dedication to scientific projects.

References

1. Salas, R.: La Calidad en el desarrollo profesional: avances y desafíos. Revista Cubana Médica Superior **14**(2), 136–147 (2000)
2. Abramo, G., D'Angelo, C.: Evaluating university research: Same performance indicator,different rankings. J. Informetrics **9**(3) (2015). https://doi.org/10.1016/j.joi.2015.04.001
3. de Castro Reis, J.d.: Avaliação da produção acadêmica de docentes dos cursos de pós-graduação de engenharia da Universidade Federal Fluminense utilizando o modelo DEA e índice h. Revista Meta: Avaliação, **9**(27) (2017). https://doi.org/10.22347/2175-2753v9i27.1272
4. Salomón, Y.P., Rodríguez, A.M.: Producción científica. Ciencias de la Información, vol. 38 (2007)
5. Paez, J.: La Investigación universitaria y la formación del profesorado Latinoamericano. Rev. Ciencias Sociales (2010). ISSN 1696-7348

6. Colcencias: Política Nacional para mejorar el impacto de las publicaciones científicas nacionales. Departamento administrativo de ciencia, tecnología e innovación- Colciencias, Dirección de Fomento a la Investigación (2016)
7. Barra, A.M.: La Importancia de la Productividad Científica en la Acreditación Institucional de Universidades Chilenas (2018)
8. Opazo, V.S., Soto, L., Martín, R.S.: Universidades de excelencia en Chile, según su eficiencia productiva en publicaciones científicas. Minutas de OPPES (2023)
9. Buela-Casal, G., Quevedo-Blasco, R., Guillén-Riquelme, A.: Ranking 2013 de investigación de las universidades públicas españolas. **27**(4), 317–326 (2015). https://doi.org/10.7334/psicothema2015.140
10. Mouton, J., Marais, H.C., Prinsloo, K.P., Rhoodie, N.J., Mauer, K.F.: Basic Concepts in the Methodology of the Social Sciences. HSRC Publishers (1990)
11. Salimi, N.: Quality assessment of scientific outputs using the BWM. Scientometrics **112**, 195–213 (2017)
12. Mårtensson, P.F.: Evaluating research: a multidisciplinary approach to assessing research practice. Res. Policy, 593–603 (2016)
13. Gómez-Morales, J.: Una propuesta centrada en el uso de revisiones sistemáticas para evaluar la calidad científica. Cultura Científica (2020)
14. Gelfert, A.: Critical Introduction to Testimony. Bloomsbury (2014)
15. Rosenbloom, J.L.: The effects of research & development funding on scientific productivity. PLOS ONE (2015)
16. Benavente, J.G.: The impact of national research funds: a regression discontinuity approach. Res. Policy, 1461–1475 (2012)
17. BCN: Biblioteca del Congreso Nacional de Chile, June 2021. https://obtienearchivo.bcn.cl/obtienearchivo?id=repositorio/10221/32223/2/BCN_Centros_Regionales.pdf
18. GOB: Gobierno de Chile, 16 January 2023. https://www.gob.cl/noticias/gobierno-presento-las-principales-medidas-del-plan-de-reactivacion-educativa-2023/
19. Matus, N.: Experiencia del Estudiante de la Escuela de Ingeniería Informática de la PUCV (2020)
20. Rosenbaum, M.O.: How to create a realistic customer journey map (2016)

Virtual Reality Meets Social Media: Transforming Skill Acquisition in Physiotherapy, Veterinary Surgery, and Driver Training

Sutasinee Thovuttikul[1(✉)], Chinawat Isradisaikul[2], Siriporn Peansukmanee[3], Siriphan Kongsawasdi[4], and Kittichai Wantanajittikul[5]

[1] Department of Computer Science, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand
sutasinee.th@cmu.ac.th
[2] Department of Computer Engineering, Faculty of Engineer, Chiang Mai University, Chiang Mai, Thailand
[3] Department of Companion Animals and Wildlife, Faculty of Veterinary Medicine, Chiangmai University, Chiang Mai, Thailand
[4] Department of Physical Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Thailand
[5] Department of Radiologic Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Thailand

Abstract. This study investigates the effectiveness of virtual reality technology in educational and training environments, specifically examining the impact of 360-degree video and different viewpoint perspectives on learning outcomes. Using Kolb's experiential learning theory as a framework, the research explored three case studies across diverse domains: physiotherapy training, veterinary surgery training, and new driver training. The study compared the efficacy of 360-degree video against interactive 3D simulations and analyzed the impact of first-person, second-person, and third-person perspectives on learning outcomes. The implementation utilized Oculus Quest 2 headsets, Unity 3D software, and 360Instra Version 2 cameras to create immersive learning environments. Results demonstrated that 360-degree videos provided effective learning experiences while requiring fewer resources compared to full 3D simulations. First-person perspectives proved most effective for practical skill development and patient communication (89.36% preference), while third-person views enhanced movement comprehension and reduced visual discomfort (76.60% preference). The findings suggest that a multi-perspective approach, combining different viewpoints based on specific learning objectives, yields optimal results in VR-based education. This research contributes to understanding how VR technology can be effectively implemented in professional training programs while considering both technical and pedagogical aspects.

Keywords: Virtual Reality Education · 360-degree Immersive Video · Point-of-View Learning · Experiential Learning · Professional Skills Training

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 290–302, 2025.
https://doi.org/10.1007/978-3-031-93539-8_19

1 Introduction

The rapid advancement of virtual reality technology has created new opportunities in education and training across various domains [1, 2]. This study leverages Kolb's experiential learning theory, enabling students to learn through virtual demonstrations and abstract conceptualization. In virtual environments, learners can safely experiment with different approaches and gain concrete experience through simulated real-world tasks, followed by reflective observation of their performance [3].

Virtual reality, particularly through 360-degree video and interactive 3D simulation, has become increasingly popular for practical skill development [4–6]. While comprehensive VR systems can be complex, simpler approaches like 360-degree video can provide significant benefits when properly designed for specific learning applications.

The study examines how different perspectives - first-person, second-person, and third-person viewpoints - influence learning outcomes [7]. Each perspective offers unique advantages: first-person viewpoints convey real-world experiences, while second and third-person perspectives provide alternative vantage points for comprehensive understanding [8].

Our research applies these perspectives across three domains: physiotherapy training, veterinary surgery training, and new driver training. The methodology includes comparative experiments between first and third-person perspectives in physiotherapy, role-based viewpoints in veterinary surgery, and self-reflection through second-person observation in driver training.

The subsequent sections cover related work on VR in learning (Sect. 2), system description and hypothesis (Sect. 3), experimental settings and results (Sect. 4), and conclusions with future research directions (Sect. 5).

2 Literature Review

Virtual reality can create realistic immersive videos and interactive 3D simulations that can provide learners with a safe and controlled environment to practice and develop their skills without the risks associated with real-world scenarios. For instance, in the medical field, VR has been utilized to train healthcare professionals, such as physicians and nurses, in various procedures, including surgical techniques, patient care, and emergency response scenarios. [4, 5, 6].

2.1 360-Degree Videos vs Interactive 3D Simulation System

360-degree video technology provides an immersive learning experience by allowing learners to actively explore virtual environments rather than passively watching traditional videos [9–11]. These videos are created using specialized multi-lens cameras and stitching software to produce a complete spherical view [12]. While offering immersive viewing, they typically limit interaction to looking around the pre-recorded environment [13].

In contrast, interactive 3D simulations offer more dynamic engagement, built using platforms like Unity or Unreal Engine [14, 15]. These simulations enable users to

manipulate virtual environments and receive immediate feedback [16]. They incorporate game engines, physics-based modeling, and user input devices, allowing for complex procedure practice and personalized guidance [17].

The choice between these technologies involves several considerations. 360-degree videos require lower initial costs and fewer resources, mainly needing camera equipment and stitching software [18]. However, interactive 3D simulations, despite higher upfront costs and longer development times [19, 20], offer greater flexibility for content modification.

Regarding functionality, 360-degree videos limit user control to viewing angle adjustments, while 3D simulations enable comprehensive interaction and environmental manipulation. 360-degree videos remain fixed after production, whereas 3D simulations can be continuously updated [21–23]. This makes 360-degree videos suitable for quick deployment and fixed content, while interactive 3D simulations better serve applications requiring high user engagement and complex interactions.

2.2 Perspectives in Virtual Reality: Comparing First, Second, and Third Person Views for User Experience

Virtual Reality's effectiveness heavily depends on perspective choice, which fundamentally shapes user engagement and experience. Each viewpoint offers distinct advantages and challenges that must be evaluated based on the intended application and user experience goals [9, 14, 24].

First-person perspective provides the most immersive experience, placing users directly in the virtual environment as if viewing through their own eyes. While this creates a strong sense of presence ideal for training simulations and virtual tours, it can lead to motion sickness and spatial awareness challenges during close interactions. Second-person perspective allows users to experience the virtual world through another character's viewpoint while being directly addressed. This creates an engaging psychological dynamic useful for storytelling and educational content, though it can limit user agency and pose implementation challenges. Third-person perspective positions the camera at a distance, offering improved spatial awareness and reduced motion sickness due to the stable, wider view. This makes it particularly effective for complex navigation tasks and environmental understanding, though it may reduce immersion compared to first-person experiences. The choice of perspective should be guided by the specific learning objectives and user requirements of the VR application.

Research Questions. This study aimed to investigate two key research questions that were critical to understanding the potential of virtual reality in enhancing learning and skill development across different domains.

1. How does learning professional skills via 360-degree video help students gain knowledge, build confidence, and overcome their fear? [1–3]
2. What is a suitable point of view (first-person, second-person or third-person point of view) for optimal learning and skill development in a VR environment? [4]

3 Methodology

To address these research questions from previous section, we reviewed the existing literature on the use of immersive virtual reality in educational and training contexts, focusing on the pre-design, design, and evaluation phases, which encompass the measurement of outcomes and user experiences [4].

First, in the pre-design, we need to identify the purpose of the VR system and the learning objective, identifying target audience, specify the learning outcome, select suitable technology, application domain, and create assessment tools. [2, 9, 25]. The key objectives for a VR system are interaction, immersion, user involvement, and photorealism. These are important factors to consider when deciding on the appropriate solution (filming 360-degree videos or creating interactive 3D simulations).

Second, in the design, we started developing learning materials, designing immersive scenes, and implementing techniques such as filming 360-degree videos and creating interactive 3D simulations that course from costs of production, technological development, immersion, and learning potential [4, 25, 26]. Points of view (first-person, second-person, or third-person) are incorporated into the design for comparison. First-person provides the most immersive experience but can lead to motion sickness. Second-person creates unique psychological dynamics useful for storytelling. Third-person offers improved spatial awareness and reduced motion sickness.

In the evaluation stage, the researchers assessed the effectiveness of the VR system in achieving the intended learning outcomes, which included knowledge acquisition, usability, skill development, user satisfaction, and confidence based on user experience [4]. The study emphasized the significance of employing multiple evaluation techniques, recommending the use of a combination of methods, such as questionnaires, interviews, data recording, and direct user observation, to comprehensively evaluate the system's performance [27].

The study conducted three case studies: virtual physiotherapy and rehabilitation training, virtual veterinary surgery training, and new driver virtual training. These case studies offer insights into how virtual reality can enhance learning and skill development. These insights can guide the design and implementation of an integrated virtual reality system that can be broadly applied to improve learning, as well as professional and communication skill development.

Case Study 1: Virtual physiotherapy and rehabilitation training. In the **pre-design phase**, the purpose of the virtual reality system was to provide physiotherapy and rehabilitation practical training for students aspiring to become physiotherapy professionals. The students would experience these scenarios from the therapist's perspective, gaining a deeper understanding of the techniques and communication strategies involved, particularly using eye contact[6].

In the **design phase**, the team decided to use 360-degree videos filmed at a physiotherapy clinic. This solution was chosen due to limitations in time and budget, while still considering the purpose of the physiotherapy and rehabilitation training. The researchers found that simply observing the scenarios, without making decisions, was sufficient for students to learn the practical aspects of physiotherapy and rehabilitation. Furthermore, the realism of the 360-degree videos featuring real patients was deemed to provide more

engaging emotional contact than the use of 3D character models. These videos captured patient-therapist interactions and exercises from the first-person perspective to immerse students in the scenarios. Furthermore, a human-like agent was incorporated to assist learners in enhancing their confidence and engagement.

During the **evaluation phase**, the researchers assessed the learning outcomes, user comfort, and learner confidence using a combination of questionnaires and interviews. Participants were interviewed about their experiences and comfort levels when using a head-mounted display during the learning simulation. Additionally, data was collected through pre- and post-simulation questionnaires to evaluate the students' confidence levels and assess the impact of different perspectives on their knowledge of the learning objectives.

Case Study 2: Virtual veterinary surgery training. In the **pre-design phase**, the purpose of the virtual reality system was to train veterinary students by immersing them in realistic surgical scenarios. Students were assigned various roles, such as surgeon assistants, anesthesiologists, and nurses, each with distinct positions and required skills within the operating room. The veterinary trainers aimed to create 360-degree videos that realistically depict animal anatomy, including blood and body parts, to reduce student anxiety and build their confidence.

The **design phase** for the virtual veterinary surgery training system allowed veterinary students to observe surgical procedures from three distinct roles: surgeon assistants, anesthesiologists, and nurses[28]. Each role had a unique perspective and positioning within the operating room. Surgeon assistants stood beside the surgeon and did not need to move extensively. Anesthesiologists positioned themselves at the upper body of the animal patient, sometimes needing to monitor devices at the animal's head and observe the entire body. Nurses had a more flexible position, assisting the surgeon or anesthesiologists as needed, and observing the overall operating room to quickly respond to various situations. In this study, the researchers assigned the first-person perspective to the surgeons and the third-person perspective to the anesthesiologists and nurses to enhance their respective roles and skill development.

In the **evaluation phase**, the researchers assessed the learning outcomes, user comfort, and learner confidence through interviews. The study examined how the different point-of-view setups based on the roles in the surgery room influenced knowledge acquisition, confidence, and fear reduction in veterinary practice. Data was collected via pre- and post-simulation interviews, as well as observation reports.

Case Study 3: New Driver Virtual Training. The **pre-design phase** of the new driver virtual training project differed from the previous case studies, as it focused on a more general issue that did not necessitate capturing the specific realism of human or animal patients. Instead, the purpose of this simulation was to promote more active learner behaviors and interactions, such as making decisions to control the car, using turn signals, and stopping the vehicle at appropriate locations.

During the **design phase**, the researchers developed a virtual reality driving simulation featuring a 3D low-polygon environment with interactive components. These interactive elements allowed learners to operate the steering wheel, accelerate, brake, and stop at appropriate location, thereby providing a comprehensive driving experience. The simulation incorporated a variety of challenging road conditions, traffic scenarios, and

intersection situations that align with real-world driving experiences. The researchers compared the effectiveness of first-person and second-person perspectives in helping learners comprehend driving rules, build confidence, and transition to actual on-road driving.

The **evaluation phase** assessed the impact of different perspective options on the learners' ability to develop essential driving skills, such as making appropriate decisions, following traffic signals, and maintaining vehicle control. Participants first experienced the simulation from a first-person perspective, then switched to a second-person view, with the camera positioned behind pedestrians walking towards a crosswalk. Data was collected through post-simulation surveys, interviews, and observations to determine the effectiveness of each perspective in promoting learner engagement, skill development, and confidence in real-world driving scenarios.

4 Experiment and Result

The results from the three case studies provided valuable insights into the effectiveness of using immersive virtual reality technologies for various educational and training applications. These case studies examined the use of VR in physiotherapy and rehabilitation training, veterinary surgery training, and new driver training, highlighting the potential benefits and considerations for implementing such technologies in different learning contexts.

The virtual reality devices used in these 3 projects were the Oculus Quest 2 headsets. The software used to build the virtual environments was Unity 3D, and the cameras used to record the 360-degree videos were 360Instra Version 2 cameras.

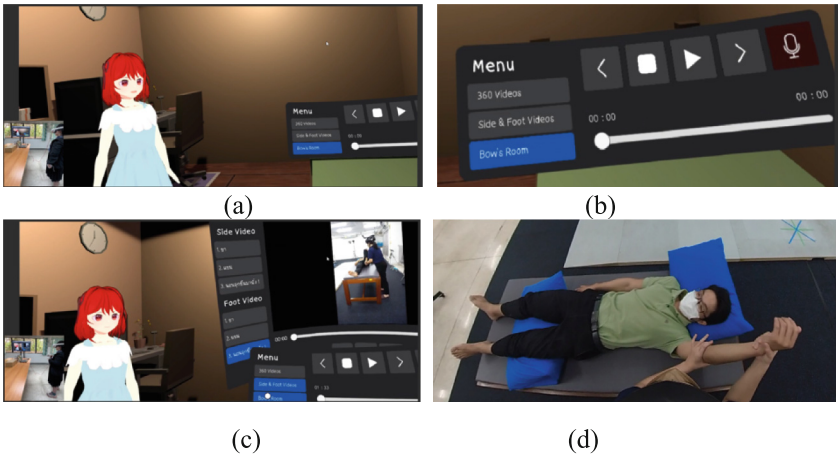
Case Study 1: Virtual physiotherapy and rehabilitation training. In the physiotherapy and rehabilitation training case study, we explored using first- and third-person perspectives in a VR environment to help students learn practical skills as see in Fig. 1.

(a) (b)

Fig. 1. Experiment setting during the participant observe in (a) first-person POV camera and (b) third-person POV camera.

We designed a **VR learning system** with the following features:

1. *Agent Introduction: An agent introduces the learning objectives and provides background information on physiotherapy and rehabilitation topics, such as range of motion and mobility exercises (see Fig. 2).*
2. *Menu Selection: Learners can choose to view physiotherapy and rehabilitation videos in either first-person or third-person perspective (see Fig. 2).*



(a) (b)

(c) (d)

Fig. 2. System setting and a Screen shot of the system (a) the agent, (b) menu, (c) third-person view, (d) third-person view.

A total of 47 students enrolled in the bachelor's degree program in physiotherapy and rehabilitation at the Faculty of Associated Medical Sciences, Chiang Mai University, were recruited to utilize the virtual reality system as shown in Fig. 2. These participants leveraged the system to acquire foundational physiotherapy and rehabilitation techniques prior to engaging in hands-on practice within their regular academic curriculum. The tasks included:

1. *Passive Range of Motiona. Arm exercisesb. Leg exercises*
2. *Functional Mobilitya. Lying down to sitting upb. Sitting to standing upc. Standing to sitting down*

After using the system from both perspectives, participants completed a questionnaire about their experience with the Physical Therapy Practice poses. The key question was whether the first-person view from an expert's perspective helped students learn physiotherapy and rehabilitation more effectively than the third-person view. The questionnaire assessed:

1. *How closely the experience mirrored real-world practice.*
3. *Insights gained about patient communication.*
4. *Clarity of the demonstrated movements.*
5. *Visual comfort throughout the learning process.*

The first-person POV was predominantly selected for its effectiveness in simulating a realistic practice experience (89.36%) and facilitating insights into patient communication (85.11%). Conversely, the third-person POV was preferred for its ability to enhance the clarity of demonstrated movements (57.45%) and to provide superior visual comfort during the learning process (76.60%).

The findings indicated that the use of first-person perspective videos and human-like agents was effective in enhancing students' confidence in communicating with patients, as it provided them with an immersive experience that simulated the perspective of the healthcare provider. Conversely, the third-person perspective videos allowed participants to feel more comfortable and relaxed, enabling them to observe the overall postures and movements involved in the Physical Therapy Practice poses.

Case Study 2: Virtual veterinary surgery training. The second case study explored using virtual reality to train veterinary students in surgical procedures on a virtual animal patient. The VR simulation captured real horse surgery procedures at the large animal hospital using 180 and 360-degree cameras. The educational materials encompassed a comprehensive series of recordings: a 10-min operation in the anesthesia room from a third-person perspective, followed by a 5-min recording of the animal's transfer to the surgery room, also from third-person view. This was complemented by a 20-min recording of operations in the surgery room from third-person perspective, and concluded with a 20-min horse castration surgery filmed from the veterinary surgeon's first-person perspective. This sequential documentation provided students with a complete view of the surgical process from multiple angles and perspectives (Fig. 3).

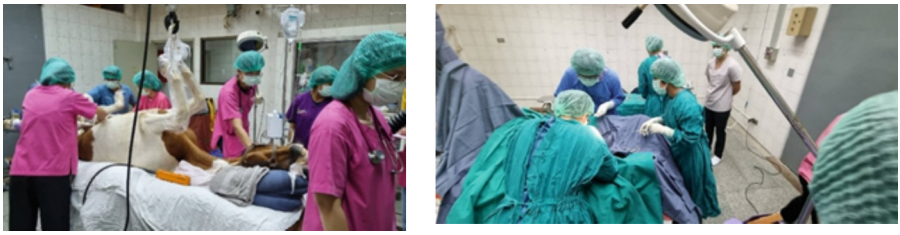


Fig. 3. Examples of horse surgery instruction showing the roles of veterinary anesthesiologist, veterinary nurse (left) and veterinary surgeon (right).

A total of 65 veterinary students from Chiang Mai University in Thailand used the virtual reality system for their Equine Practice-Clerkship course. Before and after using the VR materials, students were asked to rate their confidence, knowledge, and anxiety regarding their specific roles in large animal surgery.

The students were assigned to different roles - anesthesiologists, surgeons assistance, or nurses - and given corresponding learning materials to study. However, they could access additional materials beyond their assigned roles.

After using the VR system, data was collected on their experience with the equipment, including visual clarity, eye comfort, ease of use, and any side effects. Factors like vision conditions and use of glasses or contact lenses were also noted (Fig. 4).

Fig. 4. Example of VR simulation usage by 6th year veterinary students.

According to surveys, the virtual reality system helped, we found that 92% of students feel more confident in their duties, 86% increase their knowledge, 67% reduce their anxiety about performing their roles.

The results indicated that the majority of students experienced enhanced confidence and knowledge levels, with significant gains observed across various roles including anesthesiologists, surgeons, and nurses. Many students reported that the virtual reality system and materials were more beneficial than traditional instructional methods such as lectures, discussions, or observing real surgeries, in terms of acquiring realistic practical experience. However, a small subset of students did not exhibit any notable changes, and it was observed that some individuals assigned to the anesthesiologist and nurse roles experienced increased anxiety. Nonetheless, this did not impact their actual performance capabilities, as all students received satisfactory evaluations during their subsequent practical assessments.

Case Study 3: New Driver Virtual Training. This case study explored the use of a virtual driving simulator to assess its potential effectiveness in improving the driving skills and rule compliance of young adult participants from the Faculty of Science. The virtual environment featured urban driving scenarios with varied traffic conditions, focusing on day driving situations in a residential area.

The participants were given the following instructions: You are running late to class due to an incident, and you only have 10 min until class time. However, you need 2 min to walk from the parking lot to the classroom building, leaving you with 8 min to drive to the parking lot. The classroom building is 2 km away, and the route you need to drive through has city streets with potential cars and pedestrians, so you must drive carefully. Goal is to reach the parking lot as quickly as possible while driving safely.

In the first phase of the experiment, the study recorded the participants' driving behaviors throughout the simulation, which was presented from a first-person perspective. This included observations on whether the participants stopped at appropriate locations, their use of the brake and accelerator, their ability to maintain and vary speed, and their compliance with traffic regulations. In the second d phase, after completing the driving course, the participants had to observe their own driving behavior from the perspective of a pedestrian at a crosswalk. Afterward, the participants completed a survey with comments about their behavior. In the concluding session, they were informed that they had observed their own driving behavior (Fig. 5).

Fig. 5. Example of VR simulation at the starting point.

Fig. 6. Pedestrians POV (left) and driver POV (right).

The virtual driving simulation study identified several critical safety concerns in the interaction between drivers, pedestrians, and road infrastructure. After the participant observe their own behavior, 80% of them give a negative comment to their behaviors. In these comments 68% (17) of participant's mention driver behavior of speeding, aggressive driving with braking late, failures to yield to pedestrians, and lack of awareness of vulnerable road users. Example of the participants' comments: some drivers appear to be in a hurry and can't stop in time, difficulty judging braking distance, and some drivers assume pedestrians will wait for cars to pass first 44% (11) of participant's mention Visibility and Stopping Issues, Intersection Design Concerns, and Safety Concerns. Example of the participants' comments: difficult to see pedestrians on straight roads and stop in time, crosswalk placement is too far from the intersection, emergency situations (like rushing to hospital) may cause unsafe driving participants provided feedback on the VR driving simulation. Several users requested improvements to the vehicle feedback systems, such as better engine sound indicators and clearer signal indicators. Some found the vehicle too powerful and difficult to control, especially during turns. Visual clarity was a concern, with users reporting difficulty spotting pedestrians and seeing directional arrows. One participant suggested using the simulation for driver's license screening. There were also comments about motion-related issues, such as forward head motion during braking and visual distortions when viewing pedestrians (Fig. 6).

After reviewing videos of their VR driving performance, participants rated themselves on a scale from -2 to + 2. None of the 20 participants rated themselves as excellent drivers (+2), while 35% considered themselves above average (+1). Only 5% rated their driving as average (0), and a significant portion acknowledged room for improvement,

with 45% rating themselves as needing some improvement (-1) and 15% rating their performance as poor (-2). The average score of -0.4 indicates that participants generally viewed their driving skills as slightly below satisfactory.

Overall, the VR driving simulation provided valuable insights into the behavior of young, inexperienced drivers and highlighted areas for improvement in driver education and road infrastructure design.

5 Discussion and Conclusion

Impact of 360-Degree Immersive Video on Learning Outcomes. 360-degree immersive video technology enhances professional skill development through knowledge acquisition, with 86% of veterinary students reporting increased knowledge. The technology's effectiveness stems from multiple viewing angles, replay capabilities, and standardized learning experiences. Student confidence improved significantly, with 92% of veterinary students reporting increased confidence. Anxiety reduction varied, with 67% of veterinary students experiencing less anxiety, though some roles reported increased anxiety. The driving simulation showed that self-awareness led to improved risk perception.

Optimal Point of View for VR Learning. Different viewpoints showed varying effectiveness in VR learning. First-person POV was preferred by 89.36% of physiotherapy students for realistic practice and 85.11% for patient communication. Third-person POV proved better for movement understanding (57.45%) and visual comfort (76.60%). The research indicates best results come from combining both perspectives, with first-person POV excelling in procedural learning and third-person POV in spatial understanding.

360-degree immersive video enhances professional skill development through comprehensive learning experiences. Different learning objectives require specific viewing perspectives, suggesting VR systems should incorporate both first and third-person views. Future implementations should consider role-specific anxiety, self-reflection capabilities, and user comfort. Further research should examine long-term skill retention and optimal timing for perspective switching.

Acknowledgments. I would like to express my sincere gratitude to the Teaching and Learning Innovation Center, Chiang Mai University and Faculty of Science, for their generous support through the "Learning Innovation Research Level 1" scholarship and "Research Grant from Annual Revenue Budget for Fiscal Year 2022 in the category of Basic Research Project in Science". These supports has been invaluable in enabling me to conduct this research and advance my studies in physiotherapy and rehabilitation. I deeply appreciate the opportunities and resources provided by the center, which have greatly contributed to the success of this project.

References

1. Samadbeik, M., Yaaghobi, D., Bastani, P., Abhari, S., Rezaee, R., Garavand, A.: The applications of virtual reality technology in medical groups teaching, (2018). https://doaj.org/article/d874d35a062c4a899ef247f3d58bde99
2. Siegle, R.F., Roscoe, R.D., Schroeder, N.L., Craig, S.D.: Immersive Learning Environments at Scale: Constraints and Opportunities (2020). https://doi.org/10.1177/1071181320641278
3. Davitadze, M., Ooi, E., Ng, C.Y., Zhou, D., Thomas, L., Hanania, T., Blaggan, P., Evans, N., Chen, W., Melson, E., Arlt, W., Kempegowda, P.: SIMBA: using Kolb's learning theory in simulation-based learning to improve participants' confidence, (2022). https://doi.org/10.1186/s12909-022-03176-2
4. Checa, D., Bustillo, A.: A review of immersive virtual reality serious games to enhance learning and training (2019). https://doi.org/10.1007/s11042-019-08348-9
5. Ziegler, C., Παπαγεωργίου, A, Hirschi, M., Genovese, R., Christ, O.: Training in Immersive Virtual Reality: A Short Review of Presumptions and the Contextual Interference Effect (2020). https://doi.org/10.1007/978-3-030-44267-5_49
6. Bowyer, M.W., Streete, K.A., Muniz, G., Liu, A.V.: Immersive Virtual Environments for Medical Training (2008). https://doi.org/10.1053/j.scrs.2008.02.005
7. Gorisse, G., Christmann, O., Amato, É.A., Richir, S.: First- and Third-Person Perspectives in Immersive Virtual Environments: Presence and Performance Analysis of Embodied Users (2017). https://doi.org/10.3389/frobt.2017.00033
8. Moore, C., Barresi, J.: The Role of Second-Person Information in the Development of Social Understanding (2017). https://doi.org/10.3389/fpsyg.2017.01667
9. Mayer, R.E.: The Cambridge handbook of multimedia learning (2006). https://doi.org/10.5860/choice.43-5643
10. Antoniou, P.E., Economou, D., Athanasiou, A., Tsoulfas, G.: Editorial: Immersive Media in Connected Health (2021). https://doi.org/10.3389/fdgth.2021.697336
11. Lindsay, R., Spittle, S., Spittle, M.: Skill adaption in sport and movement: Practice design considerations for 360° VR (2023). https://doi.org/10.3389/fpsyg.2023.1124530
12. Fan, C.-L., Lo, W.-C., Pai, Y.-T., Hsu, C.-H.: A Survey on 360° Video Streaming, https://doi.org/10.1145/3329119, (2019)
13. Argyriou, L., Economou, D., Bouki, V.: Design methodology for 360° immersive video applications: the case study of a cultural heritage virtual tour, (2020). https://doi.org/10.1007/s00779-020-01373-8
14. Sveistrup, H.: Motor rehabilitation using virtual reality (2004). https://doi.org/10.1186/1743-0003-1-10
15. Leung, T.I., Zulkernine, F., Isah, H.: The use of Virtual Reality in Enhancing Interdisciplinary Research and Education (2018). https://arxiv.org/abs/1809.08585
16. Ρούσσου, M, Oliver, M., Slater, M.: The virtual playground: an educational virtual reality environment for evaluating interactivity and conceptual learning (2006). https://doi.org/10.1007/s10055-006-0035-5
17. Lean, J., Moizer, J., Derham, C., Strachan, L., Bhuiyan, Z.: Real World Learning: Simulation and Gaming. In: Springer eBooks. p. 187. Springer Nature (2020)
18. Castillo, S.A., Calvitti, K., Shoup, J., Rice, M., Lubbock, H., Oliver, K.H.: Production Processes for Creating Educational Videos (2021). https://doi.org/10.1187/cbe.20-06-0120
19. Balzerkiewitz, H.-P., Stechert, C.: The Evolution of Virtual Reality Towards the Usage in Early Design Phases (2020). https://doi.org/10.1017/dsd.2020.159
20. Lawson, G., Salanitri, D., Waterfield, B.: Future directions for the development of virtual reality within an automotive manufacturer (2015). https://doi.org/10.1016/j.apergo.2015.06.024

21. Caporusso, N., Ding, M., Clarke, M., Carlson, G., Bevilacqua, V., Trotta, G.F.: Analysis of the Relationship Between Content and Interaction in the Usability Design of 360o Videos. In: Advances in intelligent systems and computing. p. 593. Springer Nature (2018)
22. Cvetković, D.: Introductory Chapter: Multimedia and Interaction. In: IntechOpen eBooks. IntechOpen (2019)
23. Alcañíz, M., Alcañiz, J.E.B., Guixeres, J.: Virtual Reality in Marketing: A Framework, Review, and Research Agenda (2019). https://doi.org/10.3389/fpsyg.2019.01530
24. Karre, S.A., Mathur, N., Reddy, Y.R.: Is Virtual Reality Product Development different? (2019). https://doi.org/10.1145/3299771.3299772
25. Sato, S., Kageto, M.: The use of 360-degree videos to facilitate pre-learning and reflection on learning experiences, https://doi.org/10.1504/ijil.2020.107609
26. Drew, S.A., Awad, M.F., Armendariz, J., Gabay, B., Lachica, I.J., Hinkel-Lipsker, J.W.: The Trade-Off of Virtual Reality Training for Dart Throwing: A Facilitation of Perceptual-Motor Learning with a Detriment to Performance (2020). https://doi.org/10.3389/fspor.2020.00059
27. Kandi, V.R., Brittle, P., Castronovo, F., Gaedicke, C.: Application of a Virtual Reality Educational Game to Improve Design. Rev. Skill. (2020). https://doi.org/10.1061/9780784482889.057
28. Gasperin, B.D.M., Zanirati, T., Cavazzola, L.T.: Can Virtual Reality be as Good as Operating Room Training? Exper. Residen. Prog. General Surg. (2018). https://doi.org/10.1590/0102-672020180001e1397

Addressing Students' Absenteeism in Northeastern U.S. Schools: The Role of Technological Interventions for High School Students

Miselta W. Tita(✉), Johannes A. Badejo, and Joyram Chakraborty

Department of Computer and Information Science, Towson University, Towson, MD, USA
mtita@towson.edu

Abstract. Absenteeism poses a significant threat to student success; it disrupts academic progress. Despite numerous interventions, such as behavioral programs, mentorship, and disciplinary actions, the problem persists, highlighting the need for innovative solutions. This study examines the potential of digital interventions, particularly text messaging, to reduce absenteeism, based on surveys conducted with school personnel and parents/guardians in the northeastern United States. The first survey focused on stakeholders' perceptions of technology in addressing absenteeism, while the second assessed parents'/guardians' interest in real-time attendance notifications. Results from the surveys' participants highlighted a unanimous acknowledgment of absenteeism as a problem and a strong parental commitment to managing attendance through technological means. Participants valued immediate attendance notifications and the information on attendance charts, underscoring the potential of leveraging technology to enhance parental engagement and reduce absenteeism. These findings highlight the complexity of addressing absenteeism and suggest a need for balanced approaches that combine technological solutions with inclusive strategies to foster student engagement and well-being.

Keywords: Absenteeism · Attendance · Technological solutions · School · Education

1 Introduction

Absenteeism has long been a major cause of concern for parents, educators, school districts, and schools nationwide [16, 19, 20, 27]. Compulsory attendance laws were enacted to ensure students attend school regularly [11, 12]. Despite these laws, absenteeism persists, prompting schools to adopt various strategies to address the problem. Recent data indicates a troubling increase in absenteeism across various regions of the United States. For instance, Maryland's absenteeism rate rose from 18.8% in 2019 to 29.6% in 2022, while Florida's rate increased from 20% to 33% during the same period [6, 7, 24]. In response to this alarming data, schools have implemented solutions such as suspensions,

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 303–315, 2025.
https://doi.org/10.1007/978-3-031-93539-8_20

check-ins, and attendance awards to combat absenteeism. However, these strategies do not sustain long-term improvement and sometimes do not address the underlying issues contributing to absenteeism [4, 8, 13, 15, 24]. Technological interventions, including family portals and robocalls, are implemented. However, these approaches often fall short due to existing technological barriers. Inequities in technology access and varying levels of digital literacy among families result in some parents experiencing information overload, hindering their ability to effectively access essential resources and critical information via these portals [21]. The ongoing absenteeism problem underscores the necessity for innovative and more effective solutions. Text messaging offers a more effective means of communication due to its immediacy and accessibility [14, 17].

The impact of absenteeism extends beyond just students; it affects various stakeholders, including parents, school staff, and society. For schools, absenteeism disrupts lesson planning and delivery and graduation rates, which potentially reduce funding; teachers often design lesson plans based on a certain number of students being present [1]. Frequently absent students can lead to learning gaps, requiring teachers to adjust their lesson plans and sometimes reteach materials. Parents face legal issues, stress, and increased responsibilities. As Jessica Sprick and Tricia Berg emphasized in their book Teacher's Guide to Tackling Attendance Challenge, tackling absenteeism is not merely about improving attendance but equipping children and young people with the tools they need to thrive in their educational journey and beyond [22]. Consistent school attendance allows students to engage fully in the learning process, participate in extracurricular activities, and build social connections. Therefore, improving attendance is critical for building a student's solid academic foundation and career paths.

Given the limitations of current methods, there is a need for more effective technological platforms are needed to communicate attendance data to parents and guardians. While family portals have been used to communicate attendance data to parents, they often frustrate parents and fail to address absenteeism effectively, as can be seen in the continuously rising amount of absenteeism [25]. A more intentional and interactive technological solution is needed to engage students and their parents, providing real-time attendance updates and fostering accountability. This paper seeks to answer two research questions. What are school stakeholders' perceptions of using technology to solve students' absenteeism? Are parents/guardians interested in receiving real-time attendance information?

2 Literature Review

Having seen some schools attempting to tackle the issue of chronic absenteeism by implementing suspensions for students with unexcused absences, it is essential to analyze the potential consequences and effectiveness of these disciplinary measures. According to Britten Ekstrand [8], Christopher Kearney [15], and David and Linda Stone [23], this strategy can be counterproductive, as it alienates families from the educational communities that are trying to engage them. Though suspension may communicate to students that unexcused absences are not tolerated, imposing additional absences on a student who is already missing school can further disconnect the student from his educational environment. Rather than fostering a positive relationship with learning, this

punitive approach may lead to even more missed school days and does not encourage a positive attitude toward education [2, 26]. Instead of feeling inspired to attend school and learn, the student may associate school with punishment and rejection, which can lead to a further decline in attendance and engagement.

Aside from suspensions, other interventions, such as sending mail home and having students check in with an adult, have been used to curtail absenteeism. To prevent children from missing school days, schools send letters to parents about the importance of attendance [5]. In the letter, parents are informed that poor attendance can affect their children's academic performance and are given suggestions for improving their children's attendance [15]. These proactive measures foster a collaborative effort between schools and families to ensure students remain engaged and present in their educational journey. These methods of communication have their challenges: literacy levels, time constraint, clarity, and volume of communication [10]. Letters sent home are effective but only sometimes work to reduce absenteeism because parents might not read the letter thoroughly or take the necessary steps to address the child's absences, making the letter not a suitable means of changing a habit that has become a pattern.

While some schools use punitive measures to address absenteeism, others use rewards and awards. Students with perfect attendance within a semester or a quarter are awarded perfect attendance certificates, homework passes, field trips, attendance dances, and ice-cream socials with the principal [3]. Some schools motivate students by hosting a pizza party for classes with the best monthly attendance. Students also receive improved attendance recognitions, which motivates students who cannot attain perfect attendance to strive for perfect attendance. A middle school in Murray County, Georgia, recognizes students with perfect attendance and improved attendance by giving doughnuts and milk to classes with the best attendance and entering them in a prize drawing [3]. With these incentives and awards, students develop a sense of mutual responsibility to attend classes.

To reduce student absenteeism, many school districts have adopted technological methods of communicating with parents. As Balu et al. identified, one such approach involves using automated voice messages, commonly known as "robocalls", to inform parents about their child's absence from school [4]. The robocalls are generated by a system that automatically dials phone numbers and delivers a pre-recorded message notifying the parent of their child's non-attendance. However, the success of this method can vary. For instance, a student recounts in Hammond's report on attendance that after his mother received the robocall and instructed him to go to class, he did not care and continued missing his classes [13]. His persistent absences led to legal repercussions, with the student and the parent being called to court. Even after facing legal interventions, the student continued to miss class less frequently than before and was later summoned again to court due to his lack of improvement in attendance. This narrative suggests that while robocalls can alert parents to attendance issues, they may not always change the student's behavior. Additionally, the effectiveness of robocalls can further be undermined if parents do not take these alerts seriously, possibly dismissing them as mere annoyances. This dismissal can be widespread when the child convinces the parents of their presence in class, leading the parent to ignore the robocall warnings. These technologies have largely failed to maintain consistent parental engagement in their child's education [18].

3 Methods

3.1 Data Collection Method

The first question, "What are school stakeholders' perceptions of using technology to solve students' absenteeism?" delves into the attitude and beliefs of various school stakeholders-including administrators, educators, and parents- regarding adopting technological tools to tackle absenteeism. Understanding their perspectives is crucial for the successful implementation of any tech-based intervention. Stakeholders' buy-in can significantly influence the efficacy of any solution, as their support and active participation are essential for seamless integration into existing school systems. The second question, "Are parents/guardians interested in receiving real-time attendance information?" focuses on the preferences and interests of parents and guardians in receiving real-time updates on their child's attendance. Real-time information can empower parents to take immediate action, fostering a collaborative effort between the home and school to ensure consistent attendance. However, the success of such interventions depends on whether parents are willing and able to engage with these technological tools.

Two surveys and interviews were conducted, which gathered insights from stakeholders to identify school stakeholders' perceptions of using technology to reduce absenteeism and determine whether parents desire real-time information. The first survey was conducted as an interview, providing an understanding of school personnel and parents' perceptions of technological interventions. It involved eleven stakeholders: nine educators, one administrator, and one parent/guardian of high school students in a public school in Maryland, as shown in Table 1 below. All eleven participants completed an open-ended survey of ten questions.

The second interview sought to understand stakeholders' smartphone usage, interest in receiving real-time attendance notifications, and their likely response to such notifications. The study involved thirty-one participants: twenty-seven parents or guardians, one administrator, and three educators. The participants' demographic, which indicates the participants' role in the school system for the second survey, is shown in Table 2 below. The second survey was structured in a way that first filtered out respondents who did not check the attendance chart (Fig. 1), allowing for more accurate data collection from those who did.

To participate in the surveys, the participant had to be 18 years of age or older, a current administrator or educator in a public school in Maryland, or a parent/guardian of a student enrolled in a public school in the northeastern part of the US. Participants also had to be willing to share their experiences of students' absenteeism freely. All participants participated in an interview conducted virtually or physically based on availability. The researchers recorded the interviews, loaded them onto Microsoft Excel, and imported them to the International Business Machines (IBM) Statistical Package for Social Sciences (SPSS) for statistical analysis.

Table 1. First Survey's Demographic.

Participants				
	Frequency	Percent	Valid Percent	Cumulative Percent
Administrator	1	9.1	9.1	9.1
Educator	9	81.8	81.8	90.9
Parent and educator	1	9.1	9.1	100.0
Total	11	100	100.0	

Table 2. Second Survey's Demographic.

Participants				
	Frequency	Percent	Valid Percent	Cumulative Percent
Administrator	1	3.2	3.2	3.2
Educator	3	9.7	9.7	12.9
Parent/ guardian	27	87.1	87.1	100.0
Total	31	100.0	100.0	

4 Results

The results from the first survey showed a need for a solution, and many were not against the idea of using technology to curtail absenteeism. The second survey showed a high acceptance of real-time notifications and the perceived usefulness of attendance charts.

Participants in the first survey unanimously viewed absenteeism as a problem and shared their experiences. Participants indicated they had implemented various measures when asked about strategies they use to combat student absenteeism. As displayed in Table 4, these strategies include punitive measures such as administrative referrals, suspensions, and detentions; reward and incentive programs like attendance awards (certificates, homework passes, field trips) and recognition programs (class parties); supportive interventions such as mentorship programs and check-ins; and school and community resources like parent-teacher conferences. Seven participants responded affirmatively when asked if technology could address student absenteeism, while three disagreed. Those skeptical expressed concerns about the student's well-being, noting that "students may feel tracked and thus not desire to attend school."

The consensus among the stakeholders who responded to the survey question is that technology has the potential to address student absenteeism, primarily through enhanced engagement, improved communication, and efficient data tracking, see Table 4 above. The responses ranged from strongly endorsing technology to solve students' absenteeism to skepticism. Several participants believe that technology can significantly enhance student engagement. Asynchronous learning, which was widely adopted during the COVID-19 pandemic, is cited as a beneficial method. Tools like online tutoring and daily advisory

Table 3. Sample Attendance Chart.

Below is a sample attendance chart for a fictional high school student named John Doe, who is frequently absent. The chart represents one week of school, Monday through Friday, and covers a typical high school class schedule. Each day is broken into periods, with a simple indication of whether John Doe was present (P) or absent (A) for each class.

John Doe's Weekly Class Attendance Chart

Day / Period	1st Period	2nd Period	3rd Period	4th Period	5th Period	6th Period	7th Period
Monday	A	A	A	A	A	A	A
Tuesday	P	P	A	A	P	P	P
Wednesday	P	A	A	P	P	A	P
Thursday	A	A	A	A	A	A	A
Friday	P	P	P	P	A	A	P

Key:

- **P:** Present
- **A:** Absent

Class Schedule:

- **1st Period**: English
- **2nd Period**: Math
- **3rd Period**: History
- **4th Period**: Science
- **5th Period**: Foreign Language
- **6th Period**: Art (John Doe's favorite subject)
- **7th Period**: Physical Education

sessions facilitated through technology are potential ways to keep students more involved and committed to their education. Many participants highlighted the role of technology in improving communication between schools and parents, such as the SchoolMax platform that allows parents to monitor their child's attendance in real-time, while apps like Remind make it easier to send notifications and updates directly to parents' phones. Announcement videos and public service announcements about absenteeism and related issues are also suggested as effective measures. Another significant advantage noted by participants is data tracking, which is the ability to track and share student data more efficiently. This capability can help teachers and administrators to identify attendance patterns and intervene quickly. The ease of sharing data among educators and parents can keep all stakeholders informed and engaged in addressing absenteeism. Despite

the positive views, some participants expressed reservations about the effectiveness of technology in solving absenteeism. Concerns include the potential for students to feel monitored and tracked, which could lead to resistance. There is skepticism about over-reliance on technology, with some suggesting that it might contribute to the problem rather than solve it.

Table 4. Stakeholders' Perceptions

Participant	Q9: Can we use technology to solve student absenteeism? Can you please explain?	Q10: In your opinion, what are some of the barriers towards adopting technological solutions? Can you please explain?
1	"Yes, it will help students to be more engaged. Asynchronous learning was beneficial"	"Some barriers include equity issues, lack of Wifi, inconsistency, technology not being used as supposed to. When students share the internet at home, it makes it slower for them to complete assignments. It becomes a disadvantage if all students are not using the technology in class, language barriers"
2	"Yes, implement announcement videos about absenteeism and bullying, implement advisory every day, and have public service announcements"	"There will be a challenge with password reset and challenge. There will be challenges with password reset and challenges with students getting into account. If each student does not have a computer, though the school can use DonorsChoose to supplement what is needed. The challenge of teaching students how to do things related to technology in the first week of school"
3	"Yes, it can be done through online learning like during Covid or online tutoring"	"Are they getting quality learning, accountability, internet connection, quiet class to learn with background noise?"
4	"Yes, SchoolMax- parents can see when their child is absent. Remind app, technology would be easier to use on phone"	"Money, bus shortage makes it challenging to take attendance on time, lack of parents signing up, negative feedback- too much can overwhelm parents"
5	"Not sure, but it can be used to text parents"	"Lack of email for parents, error mistakenly making students absent"
6	"No, technology is more of the problem; we lean too much on it"	"Lack of access, distractions, lack of self-discipline and awareness, increase in anxiety, uncomfortable, students are not used to facing their fears"

(*continued*)

Table 4. (*continued*)

Participant	Q9: Can we use technology to solve student absenteeism? Can you please explain?	Q10: In your opinion, what are some of the barriers towards adopting technological solutions? Can you please explain?
7	"Yes, it can be used for keeping data"	"Lack of knowledge of resources available, block school numbers, parents lack of checking SchoolMax, parents don't attend virtual events, parents don't communicate with technology in the same way. Do not prioritize technology. Students would find ways to cheat it, technology would get faster, and convenience is not always the best option"
8	"No, using technology as an extension makes students feel tracked"	"Students not giving correct information, lack of proper technological device, lack of internet access, students need to be in school room to interact with a person, Chromebook is compromising students' honesty. There are language barriers. If a student is not proficient in reading language, they will struggle. Different cultures have different types of value; education is not valued today. Teachers should model behaviors"
9	"I don't see how other than for tracking purposes or perhaps as part of an incentive program"	"When the internet is down, we do not have access to technology. When students do not have their Chromebook due to loss, theft, or leaving them at home they do not have access to technology. If the system glitches, it poses a problem as well"
10	"Yes"	"Software needs to be developed that sends notifications to parents' phone and email regarding absences, tardies, and grades"
11	"Yes. Technology is a great tool that can certainly help with absenteeism. Students' data can be tracked more easily and shared among teachers. Parents have access to student records to stay abreast of attendance issues"	"Barriers include learning curve for those not as adept with technology, administrators who place blame and burden on teachers solely, and parents who are not involved in their child's education"

Though a technological solution to absenteeism would be beneficial, many were concerned about its barriers depending on the technology being implemented. The barriers can be broadly categorized into equity issues, technical difficulties, user engagement, and systemic inefficiencies. Regarding equity, the concern is mostly centered on the discrepancy in access to technology and the Internet, which worsens when multiple household members share limited resources. This can disadvantage students who lack consistent, high-speed internet or adequate devices. Language barriers also pose a significant obstacle, especially for families who are not proficient in the language of instruction. The technical difficulties include password reset, system glitches, and the need for robust programs for students' log-ins.

Additionally, there are concerns about system errors, such as mistakenly marking students as absent, which can further complicate the reliability of technological systems. Also, other technical difficulties relate to students not having their devices due to loss, theft, or forgetfulness, which impedes their ability to engage with technological solutions. The barriers concerning user engagement relate to parents and students and their ability to effectively use technologies. For parents, barriers include a lack of email access, reluctance to engage with school communication platforms, and not prioritizing the use of technology in their child's education. For students, issues include distractions, lack of self-discipline, and potential increases in anxiety. There is also skepticism about the authenticity of student engagement when using technology, with concerns about cheating and completion of tasks. Regarding systemic inefficiencies, there are financial burdens of implementing technological solutions, bus shortages that impact timely attendance recording, and administrative practices that may place undue blame on teachers. Additionally, the learning curve for students and teachers in adopting new technologies requires significant training and ongoing support.

Participants of the second survey had smartphones and used text messaging features on their devices. Participants were shown a sample attendance chart of a fictional high school student and were asked if they were aware of the attendance chart through the family portal. Eleven of the thirty-one participants were aware that they had access to the attendance chart through the family portal and had used it; thirteen participants were aware but had not used it; five participants were unaware, and two participants did not have any access to the family portal. While some of the participants use the attendance chart, a significant number choose not to engage with it or are unaware of its availability.

After being shown a sample attendance chart of a student named John Doe, participants were asked, "If John Doe was your child and you received notification immediately after attendance was taken in his class, would you use that notification to help him stop absenting his class?" All thirty-one participants said, "Yes." This result indicates unanimous agreement among the participants regarding the usefulness of immediate attendance notifications in helping to address student absenteeism. Participants were then asked to rate the usefulness of the information provided on the attendance chart; twenty-two participants stated it was extremely useful, eight stated it was very useful, one stated it was moderately useful, none stated it was slightly useful, and none stated it was not useful (Table 5). The result indicates a high perceived usefulness level among participants regarding the information provided on the attendance chart. 71% rated the information on the attendance chart as "extremely useful," 25.8% rated the information

as "very useful," and 3.2% rated the information as "moderately useful." This unanimity can help in garnering broader support for initiatives that involve the use of such attendance charts. The lack of neutral or negative responses emphasizes the strong positive perception of the attendance chart's usefulness.

Table 5. Perceived Usefulness of Information Found on the Attendance Chart.

Rate the usefulness of the information provided on the attendance chart				
	Frequency	Percent	Valid Percent	Cumulative Percent
Extremely Useful	22	71	71	71
Very Useful	8	25.8	25.8	96.8
Moderately Useful	1	3.2	3.2	100.0
Slightly Useful	0	0	0	
Not Useful at All	0	0	0	
Total	31	100	100	

When participants were asked if knowing that their child (children) was absent during each class was useful to them, twenty-nine of the participants said, "Yes," and two said, "No." 93.5% responded affirmatively, indicating that they believed knowing whether their child was absent during each class would be useful. 6.5% responded negatively, indicating that they did not find such detailed attendance information useful (Table 6).

Table 6. Is Real-time Class Absence Tracking Valuable to School Stakeholders?

Would knowing that your child (children) was absent during each class be useful to you?				
	Frequency	Percent	Valid Percent	Cumulative Percent
Yes	29	93.5	93.5	93.5
No	2	6.5	6.5	100
Total	31	100.0	100.0	

5 Discussions

The results of participants with access to the attendance chart are alarming. There is a potential need to improve communication about family portal features to ensure that more parents can effectively use the tool to track their children's attendance. The unanimous response of participants after being asked, "If John Doe was your child and you received notification immediately after attendance was taken in his class, would you use that notification to help him stop absenting his class?" also reflects a high parental

commitment to addressing and managing their child's attendance. Parents are willing to take action based on real-time information to ensure their child attends classes regularly. The results of participants who value real-time class absence tracking also suggest that most parents see significant value in having detailed and timely attendance information.

6 Recommendations

While technological solutions hold promise for addressing student absenteeism, all stakeholders must take a comprehensive approach to address the abovementioned barriers. This includes ensuring equitable access to resources, providing robust technical support, fostering user engagement, and addressing systemic inefficiencies. A possible way of overcoming the barriers mentioned by various stakeholders includes developing software that sends notifications to parents' phones and emails and leveraging platforms like DonorsChoose to supplement technological needs. Notification software will enable better communication and engagement strategies to ensure parents are more involved in their children's education through technology.

Parents have different preferences and needs when monitoring their child's attendance. Thus, schools might need to offer multiple methods of communication and monitoring to accommodate these varying preferences. The preliminary nature of the findings highlights the need for further research with a more extensive and diverse sample to validate and expand upon the initial insights. The next step is to design and implement a technological tool to provide real-time alerts to parents and guardians of a child's absence.

With the development of a technological tool that provides real-time alerts, school personnel will need an understanding of the tool and will benefit from professional development on effective ways to submit accurate student attendance that will be monitored and tracked by the school's attendance platform. School personnel can also advocate for local and national policies regarding the use of technology in school and policies regarding students' privacy and data security. Schools can use the tool to collect ongoing feedback from stakeholders to assess and refine the technological interventions being implemented continuously.

7 Limitations

The sample size of the first and second surveys is smaller and thus cannot be used to conclude an entire school or county. The findings might reflect the unique characteristics of the participants rather than the general population. One of the main limitations is the challenge of capturing the diversity of experiences and perspectives in the larger population. In the first survey, only one administrator and one parent participated, which is a significant lack of representativeness as the views of a single administrator and a single parent cannot capture the diverse perspectives, experiences, and opinions of the broader population of administrators and parents. In the second survey, there was only one administrator and three educators; thus, the perspectives and experiences of these groups are not known.

8 Conclusion

Absenteeism remains a significant challenge in the education system, affecting students' academic performance and future opportunities. Traditional interventions have proven inadequate, highlighting the need for innovative solutions. Parents and guardians who participated in the surveys see the need for interactive technological intervention, such as text messaging. By providing real-time updates and fostering accountability, these interventions can engage students and their families, improving attendance rates and ultimately enhancing educational outcomes.

While most participants saw the potential for technology to reduce absenteeism, a few were uncertain or skeptical. These mixed opinions indicated that while technology can be a valuable tool, its implementation must be thoughtful and balanced with other non-technological interventions to be truly effective.

The fact that most participants want real-time class attendance information suggests that parents are keen to stay informed about their child's whereabouts and school engagement. Given the strong positive response on both surveys, implementing technological solutions that provide real-time attendance updates, such as apps or automated notifications, would likely be welcomed by most parents. Real-time notification could enhance parental engagement and support efforts to reduce absenteeism.

References

1. Allen, C.W., et al.: School Absenteeism in Children and Adolescents. American Family Physician **98**(12), 738–44 (2018)
2. "A radical rethinking of schools to fight student absenteeism." *The Washington Post,* (2023). https://search.ebscohost.com/login.aspx?direct=true&db=bwh&AN=wapo.886fc82a-1c22-11ee-b3c8-224ae2593712&site=eds-live
3. Balfanz, R., Byrnes, V.: The Importance of Being in School: A Report on Absenteeism in the Nation's Public Schools: Promising Approaches. Educ. Digest **78**(2), 4–9 (2012)
4. Balu, R., et al.: Can Informing Parents Help High School Students Show Up for School?" *MDRC*, (2016). http://www.mdrc.org/sites/default/files/NewVisionsRCT_2016_Brief.pdf
5. Chang, H.N., et al.: Chronic Early Absence: What States Can Do. Phi Delta Kappan **98**(2), 22–27 (2016). https://doi.org/10.1177/0031721716671901
6. Chronic Absenteeism: *CT.Gov* (2024). https://public-edsight.ct.gov/students/chronic-absenteeism
7. Chronic Absenteeism Data: *Maryland State Department of Education*, (2023). https://reportcard.msde.maryland.gov/Graphs/#/Demographics/ChronicAbsenteeism/3/99/1/6/99/XXXX/2022
8. Ekstrand, B.: What It Takes to Keep Children in School: A Research Review. Educ. Rev. (Birmingham) **67**(4), 459–482 (2015). https://doi.org/10.1080/00131911.2015.1008406
9. Hiatt, J.S.: The Truant Problem and the Parental School. Bulletin, no. 29, 1915. United States Bureau of Education, Department of the Interior, Whole Number 656 (1915)
10. Epstein, J.L. (ed.): School, Family, and Community Partnerships: Your Handbook for Action, 4th edn., pp. 153–202. Corwin, A SAGE Company (2019)
11. First Public School in America: *National Geographic Society*, (2024). https://education.nationalgeographic.org/resource/first-public-school-america
12. Grocke, V.: Compulsory Education. (2024). https://www3.nd.edu/~rbarger/www7/compulso.html.

13. Hammond, B.: Getting Kids to School. Empty Desks: Oregon's Absenteeism Epidemic, (2014). www.oregonlive.com/absent/chapter-5-side1.Html
14. Heppen, J.B., et al.: Can Texting Parents Improve Attendance in Elementary School? A Test of an Adaptive Messaging Strategy. NCEE 2020–006. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (2020)
15. Kearney, C.A.: Managing School Absenteeism at Multiple Tiers : An Evidence-Based and Practical Guide for Professionals. Oxford University Press, UK (2016)
16. Kiani, C., et al.: Chronic Absenteeism: A Brief Review of Causes Course and Treatment. Adolescent Psych. **8**(3), 214–30 (2019). https://doi.org/10.2174/2210676608666180709155116
17. Kurki, A., et al.: How to Text Message Parents to Reduce Chronic Absence Using an Evidence Based Approach. NCEE 2022–001, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, (2021)
18. Lewin, C., Luckin, R.: Technology to Support Parental Engagement in Elementary Education: Lessons Learned from the UK. Comput. Educ. **54**(3), 749–758 (2010). https://doi.org/10.1016/j.compedu.2009.08.010
19. Nairn, J.: Chronic Absenteeism: Far-Reaching Consequences and No Easy Solutions. BU J. Grad. Stud. Educ. **14**(2), 15–24 (2022)
20. Rogers, T., Duncan, T., Wolford, T., Ternovski, J., Subramanyam, S., Reitano, A.: A Randomized Experiment Using Absenteeism Information to "Nudge" Attendance. REL 2017–252. In: Regional Educational Laboratory Mid-Atlantic (ED), ICF International (2017)
21. Selwyn, N., et al.: Providing a Platform for Parents? Exploring the Nature of Parental Engagement with School Learning Platforms: Providing a Platform for Parents? J. Comput. Assist. Learn. **27**(4), 314–323 (2011). https://doi.org/10.1111/j.1365-2729.2011.00428.x
22. Sprick, J., Berg. T.: Teacher's Guide to Tackling Attendance Challenges. ASCD, (2019)
23. Stone, D.H., Stone, L.S.: Dangerous and Disruptive or Simply Cutting Class: When Should Schools Kick Kids to the Curb? An Empirical Study of School Suspension and Due Process Rights. J. Law and Family Stud. **13**, 1–42 (2011)
24. Tracking State Trends in Chronic Absenteeism: *FutureEd*, (2024). www.future-ed.org/tracking-state-trends-in-chronic-absenteeism/
25. United States: Department of Education, issuing body, and issuing body United States. Department of Education. Chronic Absenteeism in the Nation's Schools: An Unprecedented Look at a Hidden Educational Crisis. Updated October 27, 2016., United States of America Department of Education, (2016)
26. Ward, S.F., Delessert, E.: LESS THAN ZERO: Schools Are Rethinking Zero Tolerance Policies and Questioning Whether the Discipline Is Really Effective. ABA Journal **100**(8), 54–61 (2014)
27. Young, S., et al.: Addressing Chronic Absenteeism in Middle School: A Cost-Effective Approach. Children & Schools **42**(2), 131–38 (2020). https://doi.org/10.1093/cs/cdaa009

Supporting Communication and Psychological Well-Being

Evolution of Emotional Response of PLEA—An Embodied Virtual Being with Emotional Capabilities

Gordana Jugo[1(✉)], Leon Koren[2], Ivan Peko[3], and Tomislav Stipancic[2]

[1] Croatian Academic and Research Network – CARNET, Josipa Marohnica 5, Zagreb, Croatia
gordana.jugo@carnet.hr
[2] Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, Ivana Lucica 5, Zagreb, Croatia
[3] Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, University of Split, Rudera Boskovica 32, Split, Croatia

Abstract. This paper explores the incremental development of PLEA, a humanoid intelligent virtual agent with emotional capabilities, that resides in virtual and physical environments. PLEA's evolution encompasses advanced emotion recognition and response generation. The study presents a survey that evaluates PLEA's ability to foster intuitive and engaging human-robot interactions. Findings highlight mixed user perceptions, with participants acknowledging PLEA's moderate ability to mimic human emotions but describing interactions as somewhat artificial. Significant differences in perceptions between students and professionals were noted, with the latter group feeling more relaxed during interactions. The paper discusses key challenges, such as achieving seamless anthropomorphism while avoiding the "uncanny valley" phenomenon and emphasises the need for iterative design improvements based on user feedback. By advancing PLEA's emotional intelligence, this research aims to enhance human-robot interaction, paving the way for applications in education, therapy, and customer service while addressing critical ethical considerations.

Keywords: Affective Robotics · HRI · Virtual Beings · Embodied Intelligence

1 Introduction

Bartneck and Forlizzi [1] created design guidelines for social robots more than twenty years ago. One of them was that they should replicate human-to-human dialogue in interactions with humans. Social robots should recognise and respond to all communication cues naturally used by humans, including verbal cues such as speech, tone, and intonation, and non-verbal cues like gestures, posture, and stance. Furthermore, the authors emphasise that robots should only express emotions they genuinely experience, as faked emotions could lead to negative perceptions.

Research in AI focusing on human interactions, particularly with robots or virtual agents, has significantly expanded over the past two decades to encompass concepts

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 319–334, 2025.
https://doi.org/10.1007/978-3-031-93539-8_21

related to affective processes, as noted by Kappas and Gratch [2]. Affective computing addresses challenges such as the diagnosis of users' emotional states to enhance interactions, as well as demonstration of affective behaviour towards users. This area of research is often grounded in two key assumptions: incorporating artificial emotional intelligence will enhance human-computer interaction, and our understanding of affective behaviour in human interaction is sufficient to guide artificial systems in replicating it effectively.

In the human-robot interaction (HRI) field, new horizons are opening with the development of emotionally intelligent robots capable of intuitive, authentic, and natural interactions. To achieve this, it is necessary to create robots that can observe, interpret, and express emotions [3].

This paper focuses on PLEA, a humanoid intelligent virtual agent with emotional capabilities, operating in both virtual and physical environments. The paper describes the evolution of PLEA's emotion recognition and response generation. It presents a survey designed to assess whether PLEA's development is progressing in the right direction while providing valuable insights for its continued improvement. An instrument for evaluating human perceptions and attitudes toward interactions with PLEA was developed and used for the first time; however, it requires further refinement. The paper comprises the following sections: a literature review on affective human-robot interaction, an overview of the incremental development of PLEA's emotional capabilities, the methodology of the conducted survey, its results, a discussion, and a conclusion.

2 Literature Review

Affective human-robot interaction has been extensively studied, as demonstrated in the following paragraphs.

Current research adopts a multimodal approach in emotion recognition systems [4–6]. Marcos-Pablos et al. [6] highlight the advantages of multimodal approaches over unimodal ones. However, emotion recognition remains a challenge for robots because it requires reliable results to ensure trustworthy interactions, and the ability to account for the recognised emotion within the time constraints necessary to adapt the robot's behaviour accordingly [3].

According to Mishra et al. [7], robots capable of emotional responsiveness can adjust their behaviour and reactions based on a user's emotional state, resulting in more natural and fluid interactions between humans and robots. Further, emotionally intelligent robots hold the potential to improve user experience, promote effective communication, and build stronger connections with humans. Spezialetti et al. [3] argue that robots that can recognise and interpret human emotions would be more effective in interacting with people. However, Mishra et al. [7] emphasise that accurately modelling emotions in robots remains a challenging and evolving area of research because emotions are intricate, multidimensional phenomena that encompass a blend of physiological, cognitive, and expressive components.

To facilitate smooth and natural human-robot interaction, robots must align with human social norms, including moods and emotions as integral aspects of social interaction. These expressions convey messages like "I'm happy to see you" or "I want to be comforted". Shared emotional experiences are essential for building long-term relationships. Social robots, particularly those used in healthcare settings or museums, must

communicate naturally and be easily understood, even by individuals without technical expertise. One proposed way to enhance these interactions is for robots to exhibit dynamic moods and emotions, similar to humans [8].

Affective behaviours allow social robots to express their internal states and to strengthen their connections with humans. Emotions are crucial in signaling understanding within Human-Robot Interaction (HRI). Studies have shown that emotions generated by the language model were reliably recognised, and participants successfully perceived the robot's emotions. Robots displaying congruent, model-driven facial emotion expressions were significantly more human-like, emotionally appropriate, and capable of creating a more positive impression [7].

Implementation of affective robots raises substantial ethical concerns, such as issues related to privacy, data security, social and cultural biases, equity of access, and effects on human-human interaction and relationships [9–11]. The anthropomorphism of robots introduces another layer of ethical complexity.

According to Bartneck et al. [12], anthropomorphism involves attributing human form, characteristics, or behaviors to non-human entities, such as robots, computers, and animals. Highly anthropomorphic robots, however, may encounter the "uncanny valley," suggesting that as a robot's appearance and movements become more humanlike, human emotional responses become increasingly positive and empathetic—up to a certain point. Beyond this point, the response shifts abruptly to intense repulsion. Nevertheless, as the robot's appearance and movements further converge with humans, emotional responses regain positivity and approach levels of empathy seen in human-to-human interactions.

A positive perception of safety is essential for fostering acceptance of robots as partners and co-workers within human environments. Bartneck et al. [12] define perceived safety as the user's assessment of the potential danger when interacting with a robot, alongside their comfort. However, direct measurements of perceived safety and user comfort have been infrequent.

User satisfaction ultimately determines how well social robots perform. Therefore, assessing users' perceptions of social robots is essential, as these perceptions cannot be directly measured within the robots themselves [12]. A deeper understanding of people's attitudes toward robots will help guide future research, development, and deployment of robotics across various public and private sectors [13].

3 Evolution of Emotional Response of PLEA

PLEA is a humanoid intelligent virtual agent capable of operating in virtual and physical environments. In its physical embodiment, PLEA uses a humanoid robot head made of a transparent plastic shell with an integrated projection mechanism that displays the virtual agent's face on the curved surface [14, 15].

PLEA perceives and recognises human emotions, generating adapted emotional responses in its interactions with people. The emotion recognition process is based on multimodal inputs, including facial expression recognition, ambient sound levels, and the intensity of body movements [15]. PLEA utilises the Facial Action Coding System FACS [16], an exhaustive system for categorising human facial expressions, to identify emotions through facial movements.

PLEA's emotional capabilities have been developed incrementally. The development of PLEA's emotional capabilities, including emotion recognition based on facial expressions and other sources, and emotion generation, went through several stages.

In a study by Stipancic et al. [15], PLEA was a mediator between a teacher and students, presenting lectures while mimicking the teacher's facial expressions. In this Wizard of Oz (WoZ) approach, the teacher controls the robot's actions without the students' awareness [17]. While PLEA could express emotions on its face, it did not produce emotional responses in this case [15]. It was merely a reproduction of the teacher's emotions. However, PLEA generated hypotheses about the students' emotions based on visual and audio inputs from its sensors and informed the teacher about their emotional status. The teacher could use this information to adapt the presentation style and establish better rapport.

The next stage in PLEA's evolution involved mirroring human emotional states during nonverbal interactions [14]. In this stage, PLEA used visual data to retrieve information and autonomously generate emotional expressions in real time through facial animations. These animations were generated with Unreal Engine software and displayed on the robot's 3D face, which is made of transparent plastic. This setup enables the display of various emotional states, such as smiles or expressions of surprise, through specific light configurations. The study [14] highlighted PLEA's ability to replicate and respond to human facial expressions, interpreting emotions, and reflecting them through similar expressions. PLEA was programmed to mimic human expressions during interactions with humans. Each interaction resulted in a series of emotional responses from the robot, which were exchanged with the person.

Koren and Stipancic [18] describe the acoustic and linguistic capabilities of PLEA, which rely on a recurrent neural network. PLEA integrates linguistic and acoustic elements to process spoken language and capture nuances such as tone, pitch, and emotion.

In conclusion, PLEA's multimodal reasoning allows it to interpret visual and auditory cues, "understand" human behaviors and emotions, and engage in meaningful non-verbal interactions.

In the third stage of its development, PLEA gained the capacity to produce its emotional reactions. Being an affective virtual entity, PLEA functions independently, modifying its reaction in real time to the emotional states of interacting people to create more engaging, natural, and sympathetic interactions.

PLEA's computational architecture, a context-to-data interpreter, interprets social signals to generate hypotheses and produce non-verbal expressions using information visualisation techniques. By leveraging a Convolutional Neural Network (CNN), PLEA can interpret complex facial expressions in real-time and respond promptly. The visual modality algorithm combines a residual neural network for face identification and a convolutional neural network for expression analysis. The facial region is cropped, resized, and processed using a pre-trained MobileNet version 7 model [14]. The system achieves a validation accuracy of 64.7%, with the confusion matrix shown in Fig. 1.

As a virtual agent, PLEA employs a multimodal strategy; however, this research focuses solely on visual data for information retrieval.

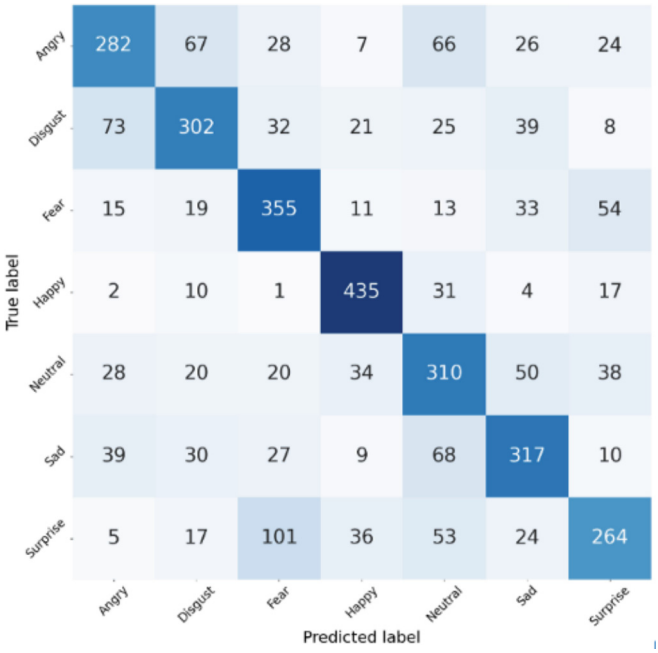


Fig. 1. Visual modality confusion matrix [14].

4 Methodology

A deeper understanding of people's attitudes toward robots can inform future research, development, and deployment of robotics across various public and private sectors [13]. This research evaluated human interactions with PLEA, recently enhanced by new emotional capabilities. The survey aimed to assess whether PLEA's development is progressing in the right direction and to provide valuable insights for its continued improvement.

The research questions were as follows:

How do participants perceive ethical and societal implications of AI-based systems like PLEA, particularly regarding learning, development, manipulation, communication, and creativity?

To what extent do participants perceive PLEA as a humanlike robot in terms of appearance and interaction?

What are the potential psychological and emotional effects of interacting with PLEA, including feelings of anxiety, relaxation, or discomfort?

The sample included 42 respondents, aged 17 to 49 (average age: 28.3), with the gender distribution of 23 men and 19 women. All participants were Croatian citizens with high levels of technological literacy. The sample included two distinctive groups: secondary school students specialising in mathematics and computer science (18 respondents) and employees of a public institution focused on digital technology and its application in education (21 respondents).

The sessions involving human interaction with PLEA were conducted in June 2024 at two different times, each held in a dedicated session room. In both sessions, PLEA was placed on a desk and prepared for interaction with the participants. Participants were called into the room individually, instructed to interact with PLEA, and then seated in front of her. They engaged in two-minute nonverbal interactions with PLEA, during which its sensory system collected visual inputs from the participants, using them for emotion recognition and subsequent emotional expression. As participants looked at PLEA and adjusted their facial expressions, PLEA responded with corresponding expressions. After the two-minute session, participants left the room. The interaction session setups are shown in Figs. 2 and 3.

Fig. 2. Interaction session setup in Room 1. PLEA is ready for interaction with participants.

Fig. 3. Interaction session setup in Room 2. PLEA is interacting with a researcher.

After the interaction, participants completed a post-interaction questionnaire to assess their perceptions of the interaction with PLEA, their attitudes toward PLEA,

the survey, and artificial intelligence in general. The questionnaire comprises 15 statements, divided into three categories: perceptions and attitudes toward PLEA (P), attitudes toward the research (R), and general attitudes toward artificial intelligence (AI), as shown in Table 1. The original questionnaire was written in Croatian, and an English translation was prepared for this paper.

Table 1. The questionnaire items and their subscales.

No	Questionnaire item	Subscale
1	Rate the research on a scale from 1 to 5	S
2	My views on artificial intelligence have changed after interacting with PLEA	AI
3	I feel that PLEA lags in interaction	P
4	I believe the conducted research is significant	AI
5	I feel uncomfortable during the interaction with PLEA	P
6	I find the interaction with PLEA artificial	P (A1)
7	PLEA's facial movements very well reflect human facial movements	P (A2)
8	PLEA's reactions surprise me	P (S1)
9	I believe that such a system could manipulate people	AI
10	I feel relaxed during the interaction with PLEA	P (S2)
11	I feel a sense of anxiety during the interaction with PLEA	P (S3)
12	I consider the interaction with PLEA natural	P (A3)
13	I believe such technology could support learning and development	AI
14	I think such a system could negatively impact communication in society	AI
15	The use of artificial intelligence negatively impacts human creativity	AI

The questionnaire comprised the following structure:

- Item 1: rated the survey overall using a 5-point scale: Unsatisfactory (1), Satisfactory (2), Good (3), Very Good (4), Excellent (5).
- Items 2 and 3: used categorical responses (Yes/No).
- Items 4 to 15: employed a 5-point Likert scale: Strongly Disagree (1) to Strongly Agree (5).

The questionnaire items measuring perceptions and attitudes toward PLEA are based on the survey by Bartneck et al. [12] who introduced the "Godspeed Questionnaires" to measure users' perception of robots. These instruments focus on five core HRI dimensions: anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety. They use 5-point semantic differential scales. Bartneck et al. [12] demonstrated the reliability and validity of these questionnaires through multiple empirical studies. The Godspeed Questionnaires provide a consistent framework for evaluating these dimensions, making them valuable tools for robot developers to monitor and refine their designs over time.

Two Godspeed Questionnaires, aligning with the research questions, were chosen to assess Anthropomorphism and Perceived Safety in this survey. The relevant items for each construct were as follows: Artificial (6), Lifelike (7), and Natural (12) for Anthropomorphism, and Surprised (8), Relaxed (10), and Anxious (11) for Perceived Safety.

This survey employed a 5-point Likert scale, following the approach used by Kulic and Croft [19]. In the presented experiment, participants rated their emotional responses to the robots in the following categories: anxiety, calmness, and surprise. The Likert scale ranged from 1 to 5, with 5 representing "extremely" or "completely" and 1 representing "not at all."

After data collection, the first step was preparing the data for analysis. Numerical values were assigned to categorical responses ("Yes" = 1, "No" = 0) and the groups of respondents: students were coded as Group = 1, while employees were coded as Group = 0 in the dataset. For negatively worded items, the scale was reversed to ensure consistency, such that lower scores reflected negative attitudes, and higher scores indicated positive attitudes.

The next step involved calculating the frequency distribution of responses for each item, the mean and the standard deviation. A reliability test was conducted before calculating the mean scores for Anthropomorphism and Perceived Safety, resulting in Cronbach's Alpha values for each category.

Cronbach's Alpha for Anthropomorphism was 0.785, indicating acceptable reliability. However, Cronbach's Alpha for Perceived Safety was initially low, suggesting inadequate reliability. To address this, item 8 was removed, which significantly improved the reliability of the Perceived Safety category. The updated Cronbach's Alpha for Perceived Safety was 0.712, indicating acceptable reliability.

5 Results

The following section presents the questionnaire results.

Table 2 displays the frequency distribution of responses for items 1 and 4 to 15, based on a 5-point scale.

Table 2. Frequency distribution of responses on a 5-point scale.

Responses	1	4	5	6	7	8	9	10	11	12	13	14	15
Strongly disagree	1	1	3	13	3	4	3	5	0	11	4	5	5
Disagree	2	1	9	13	10	10	5	8	4	20	3	8	8
Neutral	9	11	6	15	16	13	14	9	11	7	8	10	7
Agree	18	15	15	0	11	11	11	11	11	3	15	12	14
Strongly agree	12	14	9	1	2	4	9	9	16	1	12	7	8

Table 3 presents the frequency distribution of responses for items 1 and 2, using categorical responses ("Yes" = 1, "No" = 0).

Table 3. Frequency distribution of responses using categorical responses ("Yes" = 1, "No" = 0).

Responses	2	3
Yes	37	15
No	5	27

The mean and standard deviation for each item on the 5-point Likert scale were calculated. The results are presented in Table 4 and Fig. 4.

Table 4. The mean and standard deviation for items measured using a 5-point Likert scale.

Responses	4	5	6	7	8	9	10	11	12	13	14	15
Mean	3.95	3.43	2.12	2.98	3.02	3.43	3.26	3.93	2.12	3.67	3.19	3.29
Standard Deviation	0.96	1.25	0.94	1.00	1.14	1.17	1.33	1.02	0.97	1.24	1.27	1.31

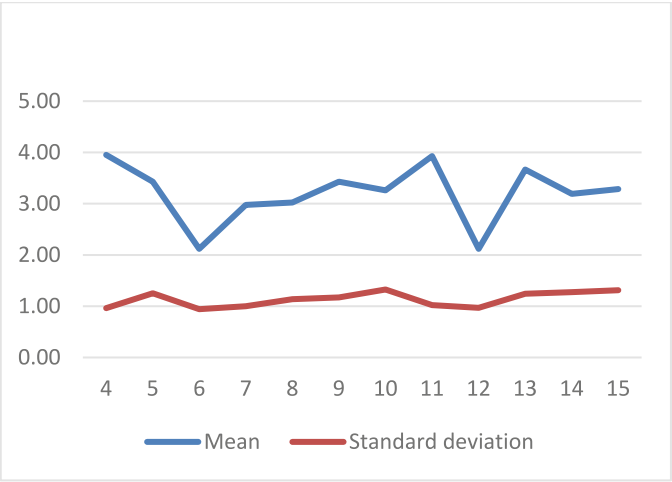


Fig. 4. Mean and standard deviation of responses 4–15 using 5-point Likert scale.

The mean and standard deviation for each questionnaire item in the Anthropomorphism category is presented in Table 5 and Fig. 5.

Similarly, the data for the Perceived Safety category are presented in Table 6 and Fig. 6.

To summarise, most answers were positive or neutral, with participants providing generally positive feedback on the research. The findings related to the interaction with PLEA revealed mixed perspectives. Respondents reported experiencing minimal anxiety during their interactions with PLEA (mean: 3.93). They also found PLEA's facial

Table 5. Mean and standard deviation of items in the Anthropomorphism category.

Item	6 (A1)	7 (A2)	12 (A3)
Mean	2.12	2.98	2.12
Standard Deviation	0.94	1.00	0.97

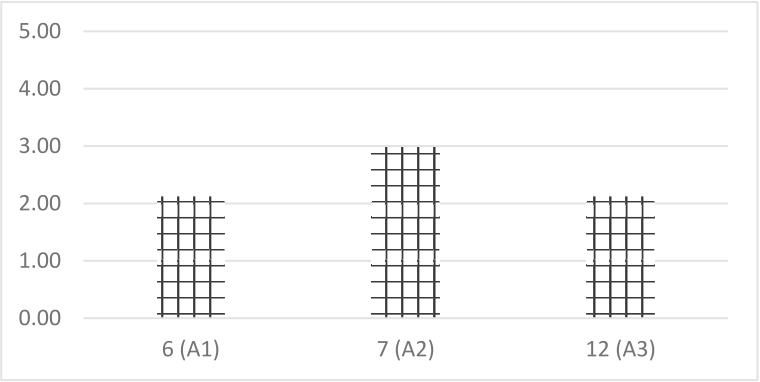


Fig. 5. Mean values of items in the Anthropomorphism category (shown in the chart).

Table 6. Mean and standard deviation of items in the Perceived Safety category.

Measure	10 (S2)	11 (S3)
Mean	3.26	3.93
Standard Deviation	1.33	1.02

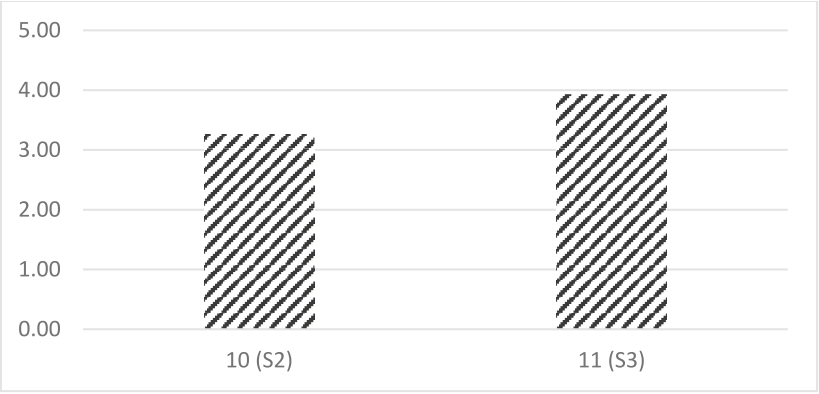


Fig. 6. Mean values of items in the Perceived Safety category (shown in the chart).

movements to resemble human expressions moderately (mean: 3.26) and reported feeling somewhat relaxed during interactions (mean: 3.26). However, participants noted that the interaction felt less natural (mean: 2.12). Additionally, participants had a generally slightly positive perception of AI's potential in learning (mean: 3.67). They expressed moderate concerns about AI's negative effects on communication and creativity (mean: 3.19), and its potential to manipulate people (mean: 3.29).

However, these results should be taken cautiously as only three items have a low standard deviation (0.94 to 0.97), indicating that the responses are clustered closely around the mean. The item related to respondents finding the interaction with PLEA artificial showed the highest agreement among participants. Very high agreement was also observed for the item with the opposite meaning, indicating respondents' attitudes toward the interaction with PLEA as natural. Respondents also had highly similar responses regarding the significance of the research. The standard deviation of other responses ranges from 1.00 to 1.33, suggesting that they are more dispersed but still show some clustering.

The means for the items related to Anthropomorphism are predominantly negative (2.12, 2.12, and 2.98), with a relatively low standard deviation. Most participants find the interaction with PLEA artificial rather than natural but are more neutral about PLEA reflecting human facial movements.

The means for items in the Perceived Safety category show that respondents are neutral about feeling relaxed during the interaction with PLEA, but most do not experience anxiety. The sense of anxiety has a lower standard deviation (1.02), suggesting that most participants agree in their answers.

To examine differences between students (Group = 1) and employees (Group = 0) in survey categories Anthropomorphism and Perceived Safety, responses were analysed by separating them based on group membership, calculating descriptive statistics, and conducting independent t-tests to compare mean responses for items A1, A2, A3 (Category Anthropomorphism) and S1, S2, S3 (Category Perceived Safety). For Category A, a significant difference was observed in item A1 (Students: $M = 1.67$, Employees: $M = 2.48$, $t = -2.94$, $p = 0.0057$), while no significant differences were found in A2 or A3. In Category S, significant difference was identified in item S2 (Students: $M = 2.78$, Employees: $M = 3.67$, $t = -2.14$, $p = 0.0389$), but no significant differences emerged for S1 or S3. These results suggest that group membership (student vs. employee) may influence responses to specific items within Categories A and S, particularly A1 and S2, while other items remain consistent across groups. This indicates that students were more prone to perceive the interaction with PLEA as artificial, while employees were less likely to share this perception. The findings suggest that employees felt more relaxed during the interaction with PLEA, whereas students were comparatively less relaxed. Overall, the results highlight differences in how the two groups experienced interactions with PLEA, with employees finding the interaction less artificial and feeling relaxed compared to students.

6 Discussion

The findings of this study offer an elaborate understanding of participants' interactions with PLEA, their perceptions of its anthropomorphic qualities, perceived safety, and their overall perception of AI. Although participants expressed concerns about AI's negative impacts on communication, creativity, and its potential for manipulation, the slightly positive overall perception of AI's potential in learning reflects optimism about its role in education and training. Participants' experiences with PLEA were mixed. They experienced minimal anxiety during interactions with PLEA and perceived its facial movements as moderately resembling human expressions. However, the interaction was described as less natural, suggesting a gap in achieving seamless human-like engagement.

The results for anthropomorphism were predominantly negative, with participants generally finding the interaction artificial but showing more neutral responses regarding PLEA's ability to reflect human facial expressions. This demonstrates a slightly more favourable perception in this specific aspect. This finding presents an opportunity to refine PLEA's design and facial expressions, enhancing its naturalness. Additionally, the low standard deviation for these items suggests that participants' perceptions were consistent, strengthening the reliability of these observations. The interaction perceived as less natural possibly indicates the presence of the Uncanny Valley phenomenon [20]. However, the study by Bartneck et al. [21] challenges the validity of the Uncanny Valley hypothesis, as the androids used in the research were equally well-liked as their human counterparts. While creating and animating androids remains a complex task, the Uncanny Valley hypothesis should no longer serve as a barrier to progress in their development.

The results for perceived safety are more encouraging. Although participants were neutral about feeling relaxed during interactions, the absence of significant anxiety demonstrates a baseline comfort with the system. This baseline comfort is an important foundation for fostering positive user experiences, as it suggests that PLEA does not provoke significant anxiety during interactions. While the neutral perception of relaxation suggests room to improve the system's ability to create a more comfortable environment, the minimal sense of anxiety is a promising starting point.

Significant differences between students and employees in some items related to anthropomorphism and safety underscore the role of group membership in shaping perceptions. Employees found the interaction with PLEA less artificial and reported feeling more relaxed, whereas students found it more artificial and felt comparatively less relaxed. These differences may arise from varying expectations, experiences or levels of exposure to similar AI systems between the groups.

The following actions could be considered to address the variability in responses. Follow-up studies could be conducted with more detailed qualitative and quantitative analyses to understand why certain items have a higher standard deviation. Another area requiring follow-up studies is the potential presence of the Uncanny Valley, which should be further explored. The variability in standard deviation also raises important considerations regarding the measurement instrument. While the high agreement on certain items suggests that the instrument reliably captures specific aspects of participants' perceptions, the more dispersed responses for other items highlight potential areas for

refinement. Further reliability and validity analysis should ensure the instrument accurately and consistently measures the intended constructs. After improving the measuring instrument, additional research should be conducted with a more robust sample, in terms of size and sampling methods. Responses should be analysed by demographic or psychographic variables (e.g., age, familiarity with robots) to identify patterns that explain the dispersion.

However, Bartneck et al. [12] propose how measurements in studies of this nature should be interpreted and highlight several inherent limitations. Determining the "ground truth" is particularly challenging, as objectively measuring how anthropomorphic a robot is can be complex. Factors such as participants' cultural backgrounds, previous experiences with robots, and their personalities may all influence the results. Conducting an experiment that accounts for all potential biases would be too elaborate and impractical. Therefore, the measurements should be seen not as absolute values, but as a comparative tool. For example, robot developers can use the questionnaires to compare different robot configurations, which can guide their decisions. However, it is important to note that human perceptions are not fixed. As people become more familiar with robots, their knowledge and expectations may evolve. Thus, the questionnaires offer only a snapshot, and if the research were repeated in ten years, the results would likely differ.

Nevertheless, these results offer important insights into how PLEA can be further developed to meet the expectations and needs of diverse user groups. The variation in standard deviations across different items indicates important factors to consider when interpreting and acting on the results. On one hand, the relatively low scores for anthropomorphism underscore the need for further development in human-like interactions to make PLEA appear less artificial. On the other hand, the lack of significant anxiety among participants offers a solid foundation for enhancing user experiences.

Bartneck et al. [12] argue that an overly anthropomorphic appearance can create expectations that the robot may not be able to meet. For instance, if a robot has a human-like face like PLEA, an uninformed user might assume it can listen and engage in conversation. To avoid potential disappointment, developers must carefully consider the level of anthropomorphism in their robot designs.

The higher standard deviations for other responses indicate a broader spread of opinions, reflecting variability in participants' experiences and perceptions. While these results still show some clustering, the dispersion suggests that certain aspects of PLEA's design or interaction experience may affect participants differently based on individual or group-specific factors, such as prior exposure to AI systems, expectations, or personal preferences.

Incorporating feedback from participants with diverse experiences addresses the wider range of user expectations. Iterative usability testing with diverse participant groups is necessary to evaluate how changes in design impact their responses. The aim is to provide a universally positive experience while addressing the unique needs of diverse user groups. This approach ensures that design improvements are targeted and inclusive, thus enhancing user satisfaction and system effectiveness.

The questionnaire is critical in developing a multidisciplinary approach to collect, analyse and interpret data from human-robot interactions involving PLEA. Once refined, it should be complemented by additional methods, such as analysis of users' facial

expressions and tone of voice, human-robot emotion-interaction flows, post-interaction interviews, and similar techniques. Combining psychological and technical aspects is crucial, as highlighted by Kappas and Gratch [2], who emphasise that research and development of artificial emotions would benefit from closer collaboration between emotion researchers and engineers.

The approach to future research and design aims to contribute to the continued development and refinement of PLEA's emotional responses to humans, fostering discussions on integrating AI into human social environments. As Kappas and Gratch [2] argue, robots that merely imitate the humans they interact with will neither replicate authentic human behaviour nor achieve long-term success. In its recent development, PLEA has gained the ability to generate emotional responses as a foundation for its advanced interactions with humans. Such human-robot interactions enhance the effectiveness of robots in roles that require emotional engagement, such as therapy, companionship, or customer service, as noted by Mishra et al. [7]. Another area of implementation is education, where several examples have shown promising results. For instance, Chew and Chua [22] integrated a humanoid robot into Chinese language learning, successfully fostering positive emotional and cognitive engagement among learners. Similarly, Aziz et al. [23] found that autistic children enjoyed the interaction sessions with robots and exhibited emotional responses, highlighting the potential of robotics in supporting special education needs.

7 Conclusion

To conclude, while participants moderately agreed that PLEA's facial movements resemble human expressions, further work is needed to refine PLEA's emotional responses. Although minimal fear was reported, the interaction was perceived as less natural, possibly indicating the presence of the Uncanny Valley phenomenon. However, it should not hinder PLEA's further progress in human-like interaction. Another area for improvement is the ongoing development of a multidisciplinary approach to collect, analyse, and interpret data from human-robot interactions, which provides valuable insights for the further enhancement of PLEA's emotional responses to humans.

While emotionally intelligent humanoid robots may offer transformative potential in therapy, companionship, customer service, and education, their integration must be guided by a robust ethical framework to ensure responsible and inclusive deployment that complements the irreplaceable role of humans.

Acknowledgements. This work was supported partially by the Croatian Science Foundation under the project "Affective Multimodal Interaction based on Constructed Robot Cognition—AMICORC (UIP-2020–02-7184)." One of the authors, Gordana Jugo, is a student in the postgraduate doctoral programme "Information Management and Application of Information and Communication Technologies," offered by the University of Mostar. Her scholarship was supported by the Croatian Academic and Research Network – CARNET. Special thanks go to Emil Grabušić and Matea Brzac, students from the Faculty of Mechanical Engineering and Naval Architecture, for their contributions to the questionnaire and research. Additional thanks are extended to the Fifteenth Gymnasium in Zagreb and CARNET for participation in this research.

Disclosure of Interests. The authors have no competing interests to declare relevant to the content of this article.

References

1. Bartneck, C., Forlizzi, J.: A design-centred framework for social human-robot interaction. In: Proceedings - IEEE International Workshop on Robot and Human Interactive Communication. pp. 591–594 (2004). https://doi.org/10.1109/roman.2004.1374827
2. Kappas, A., Gratch, J.: These Aren't The Droids You Are Looking for: Promises and Challenges for the Intersection of Affective Science and Robotics/AI (2023). https://doi.org/10.1007/s42761-023-00211-3
3. Spezialetti, M., Placidi, G., Rossi, S.: Emotion Recognition for Human-Robot Interaction: Recent Advances and Future. Perspectives (2020). https://doi.org/10.3389/frobt.2020.532279
4. Ashwin, T.S., Guddeti, R.M.R.: Automatic detection of students' affective states in classroom environment using hybrid convolutional neural networks. Educ. Inf. Technol. (Dordr) **25**, 1387–1415 (2020). https://doi.org/10.1007/s10639-019-10004-6
5. Samadiani, N., et al.: A review on automatic facial expression recognition systems assisted by multimodal sensor data (2019). https://doi.org/10.3390/s19081863
6. Marcos-Pablos, S., Alejano, F.L., García-Peñalvo, F.J.: Integrating Emotion Recognition Tools for Developing Emotionally Intelligent Agents. Int. J. Interact. Multimedia and Artif. Intell. **7**, 69–76 (2022). https://doi.org/10.9781/ijimai.2022.09.004
7. Mishra, C., Verdonschot, R., Hagoort, P., Skantze, G.: Real-time emotion generation in human-robot dialogue using large language models. Front Robot AI. **10**, (2023). https://doi.org/10.3389/frobt.2023.1271610
8. Kirby, R., Forlizzi, J., Simmons, R.: Affective social robots. Rob. Auton. Syst. **58**, 322–332 (2010). https://doi.org/10.1016/j.robot.2009.09.015
9. Ekström, S., Pareto, L.: The dual role of humanoid robots in education: As didactic tools and social actors. Educ. Inf. Technol. (Dordr). **27**, 12609–12644 (2022). https://doi.org/10.1007/s10639-022-11132-2
10. Ouafa, S., Nadir, Y., Qbadou, M., Mansouri, K.: Comparative Study of the Impact of Robot-Human Integration in the Context of Learning Scenarios: Literature Review. In: 2024 4th International Conference on Innovative Research in Applied Science, Engineering and Technology, IRASET 2024. Institute of Electrical and Electronics Engineers Inc. (2024). https://doi.org/10.1109/IRASET60544.2024.10548633
11. Newton, D.P., Newton, L.D.: Humanoid Robots as Teachers and a Proposed Code of Practice, (2019). https://www.frontiersin.org/journals/education/articles/https://doi.org/10.3389/feduc.2019.00125/full, https://doi.org/10.3389/feduc.2019.00125
12. Bartneck, C., Kulić, D., Croft, E., Zoghbi, S.: Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots, (2009). https://doi.org/10.1007/s12369-008-0001-3
13. Naneva, S., Sarda Gou, M., Webb, T.L., Prescott, T.J.: A Systematic Review of Attitudes, Anxiety, Acceptance, and Trust Towards Social Robots. Int. J. Soc. Robot. **12**, 1179–1201 (2020). https://doi.org/10.1007/s12369-020-00659-4
14. Stipancic, T., Koren, L., Rosenberg, D., Harwood, T., & Benic, J.: PLEA: The Embodied Virtual Being. In: International Conference on Human-Computer Interaction, pp. 266–275. Cham Springer Nature Switzerland. (2024). https://doi.org/10.1007/978-3-031-61281-7_18
15. Stipancic, T., Koren, L., Korade, D., Rosenberg, D.: PLEA: A social robot with teaching and interacting capabilities. J. Pac. Rim Psychol. **15**, 183449092110370 (2021). https://doi.org/10.1177/18344909211037019

16. Ekman, P., Rosenberg, E.L.: What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS). Oxford Univ. Press (2005). https://doi.org/10.1093/acprof:oso/9780195179644.001.0001
17. Riek, L.: Wizard of Oz Studies in HRI: A Systematic Review and New Reporting Guidelines. J. Hum. Robot. Interact. 119–136 (2012). https://doi.org/10.5898/jhri.1.1.riek
18. Koren, L., Stipancic, T.: Multimodal Emotion Analysis Based on Acoustic and Linguistic Features of the Voice. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 301–311. Springer Science and Business Media Deutschland GmbH (2021). https://doi.org/10.1007/978-3-030-77626-8_20
19. Kulic, D., Croft, E.: Anxiety detection during human-robot interaction. In: 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 616–621. IEEE (2005). https://doi.org/10.1109/IROS.2005.1545012
20. Rokeman, M.I., Ali, A.Z.M., Ramlie, M.K.: Talking-head, realism level and emotions in learning: A conceptual framework. Int. J. Inform. Educ. Technol. **10**, 585–589 (2020). https://doi.org/10.18178/ijiet.2020.10.8.1428
21. Bartneck, C., Kanda, T., Ishiguro, H., Hagita, N.: My robotic doppelgänger - A critical look at the Uncanny Valley. In: Proceedings - IEEE International Workshop on Robot and Human Interactive Communication. pp. 269–276 (2009). https://doi.org/10.1109/ROMAN.2009.5326351
22. Chew, E., Chua, X.N.: Robotic Chinese language tutor: personalising progress assessment and feedback or taking over your job? On the Horizon. **28**, 113–124 (2020). https://doi.org/10.1108/OTH-04-2020-0015
23. Aziz, A.A., Moganan, F.F.M., Ismail, A., Lokman, A.M.: Autistic Children's Kansei Responses Towards Humanoid-Robot as Teaching Mediator. In: Procedia Computer Science. pp. 488–493. Elsevier B.V. (2015). https://doi.org/10.1016/j.procs.2015.12.322

CLBoundaryManager: A System Facilitating Boundary Management and Border Crossing Between Life Domains

Moritz Maleck and Tom Gross(✉)

Human-Computer Interaction Group, University of Bamberg, 96045 Bamberg, Germany
hci@uni-bamberg.de

Abstract. Interruptions from non-work domains, such as family, can disrupt work tasks, particularly during periods of high cognitive load, leading to reduced efficiency and increased stress. We introduce the CLBoundaryManager concept for optimising the timing of interruptions based on relevance indications of communication requests and user workload. The implemented system uses cognitive load data from eye tracking to manage communication between work and family domains. This allows users to stay connected with family during work.

Keywords: Boundary Management · Connectedness · Interruptions · Interruption Management Systems · Cognitive Load · Eye Tracking

1 Introduction

Interruptions during work disrupt primary tasks and result in time-consuming resumption efforts and thus increase task completion time. They also affect task accuracy and efficiency [1]. This is especially the case when interruptions have a high dissimilarity to the primary task and a high complexity [2]. In offices, social codes are applied to communicate the current openness for interruptions, like opening or closing doors [3].

Such social codes are more complex to integrate when it comes to non-work interruptions, such as by other domains like family. When other family members call or message the currently working person, reasons might be diverse—asking to buy something at the supermarket before coming home, telling some irrelevant news about the neighbour, or informing about a car accident happened a few moments ago. All reasons for inquiries are of different relevance, and without prior knowledge before accepting a call from non-work, relevance assessment is nearly impossible.

Interruptions can have a positive effect, such as leading to faster completion of simple tasks or fostering creativity [1]. They strengthen social connections between individuals, including non-work social connections to strengthening family connections and foster feelings of closeness and psychological proximity [1, 4]. Yet, daily technology interruptions may also lead to negative effects if the number of interruptions is too high [5]. This highlights the need for an optimised interruption management.

Self-management of interruptions, such as disabling notifications before tasks requiring high cognitive load, may fail due to discrepancies between objective measures of

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 335–350, 2025.
https://doi.org/10.1007/978-3-031-93539-8_22

effort and subjective self-assessment [6]. Moreover, managed—compared to randomly timed—interruptions lead to better performance [7, 8].

Previous research has shown that interruptions are more harmful during phases of high cognitive load [3, 9]. Eye tracking, particularly the measurement of pupil diameter, has proven effective for assessing cognitive load—the pupil dilation does not only reflect changes in light but also cognitive processes [10]. Various systems already use the pupil diameter for real-time cognitive-load scenarios [8, 11].

Cognitive load has already been used in systems to manage interruptions, optimising the timing of interruptions to minimise disruption. [3] used signals at doors to indicate whether someone is interruptible or not, by inferring cognitive load using consumer wearables. [8] identifies low-workload moments by analysing the pupil diameter of users and tested their method in a single user scenario.

In this paper, we describe the architecture and functionality of the CLBoundary-Manager, with its two core components: a work and family user interface for managing communication requests, and the communication gateway for handling the requests based on cognitive load of the working user. The key contributions of this work are:

- A novel conept—the CLBoundaryManager—for managing communication requests and thus optimising interruption management between family and work domain.
- An implemented system that makes informed decisions based on the cognitive load derived from real-time eye tracking measurements. It does not focus on the negative effects only of interruptions and instead increases creativity by interrupting in right moments; and fosters connectedness between working and non-working family members.
- An evaluation of the CLBoundaryManager system with a user study and an in-depth discussion of the results.

To best of our knowledge we are the first to introduce an intelligent interruption management system relying on the pupil diameter to manage communication between work and family domains. The structure of this article is as follows: first, we provide a brief overview over related work. Then, we present our CLBoundaryManager concept, followed by the description of its implementation. Finally, we provide details on our evaluation of CLBoundaryManager and discuss the results and findings. We conclude our work with a conclusion and an outlook for future work. This work contributes to the development of intelligent interruption management systems to foster social connectedness.

2 Related Work

Interruptions disrupt primary tasks and often require significant time to resume them [1]. The extent of this disruption depends on factors such as the complexity of the interruption and its nature, including its similarity to the primary task [2]. The process of handling interruptions involves three main phases: first, the interruption lag, during which the user becomes aware of the secondary task; next, engaging with the secondary task itself; and finally, resuming the primary task afterward [12]. While much of the research highlights the negative impacts of interruptions, they can also have positive effects, such

as enabling faster completion of simple tasks or fostering creativity through incubation periods [1]. Yet, in work settings, interruptions often negatively impact task completion time, as well as task accuracy and efficiency [1]. This effect is particularly pronounced in team environments, where collaborative processes may be delayed, affecting not only individual performance but also the overall performance of the team [1]. Additionally, interruptions can diminish productivity and well-being among office workers [3].

Interruptions have been found to increase the cognitive load. This is mostly due to the resumption lag (that is when someone has to determine where the previous task had been interrupted to figure out the next steps) [1]. Interruptions are more harmful during phases of high cognitive load [3, 9], which gives us the design implication for interruption management systems to detect phases of low cognitive load and deliver notifications during these. Several studies exist that aim at optimising the timing of interruptions to minimise disruption by the use of the cognitive load, e.g., through the usage of workload-aligned task models [9] or measuring cognitive load through consumer wearables [3].

Social cues, such as opening or closing a door, are often used to signal openness to interruptions, while color-coded indicators (e.g., red/orange/green) can represent interruptibility [3]. However, these analogue methods become impractical without physical proximity, such as in the separation between family and work domains. For example, a family member working in an office cannot implicitly communicate interruptibility to others at home. Instead, digital solutions like switching off the phone or enabling 'Do Not Disturb' (DND) modes are commonly used. These modes often include features to define exceptions, such as allowing notifications from close family members or specific user groups or domains [13, 14].

While useful, such systems are limited—for example, they may allow all notifications from a group regardless of priority, preventing effective filtering. Additionally, users must remember to activate DND before engaging in cognitively demanding tasks, potentially reducing its effectiveness. To address this, systems often incorporate features like predefined time windows or context-aware functionalities, such as detecting a user's location, to automate interruption management [13, 15]. Extending DND functionalities could be promising, but excessive options might discourage use [14]. Furthermore, recent studies show that users fear missing critical notifications or forgetting to disable DND, adding another layer of complexity [14].

An increasement in non-work interruptions during working hours has been found by [16], specifically in the context of home-office in COVID-19. Therefore, not only an interruption management between work-from-office and family makes sense. Yet, the previously described social cues could be used, but still with the limitation of manually applying it without automated mechanisms.

Literature mostly focusses on the work-family conflict, i.e., where the family domain is disturbed by the work domain [17]. On the other hand, the family-work conflict, defined by the disruption of the work domain by the family domain, also holds negative effects. These can be on-job constructs such as lower job satisfaction or higher levels of job role conflict, and off-job constructs such as lower life satisfaction [18]. Therefore, systems should handle interruptions is in a less-disruptive way to minimise the family-work conflict. Boundary management aims at optimising the overlapping of different domains, mainly work and family domains [19].

The conflict between life domains cannot always be resolved by strictly separating them. Connectedness between life domains is often perceived very positively. Fostering connectedness (esp. with family or friends) is related to another term from the literature: 'presence-in-absence', which is to provide the sense of presence even due to physical absence [20]. Presence-in-absence further exists of themes (contact, content and context) and constituent elements (such as common ground or staying in touch for contact, personal effort for content, and unobtrusive for context) [20]. Systems to foster connectedness and presence-in-absence may derive requirements from these.

Eye tracking by the measurement of pupil diameter has shown to be a highly reliable non-invasive, real-time indicator for cognitive load, also compared to other eye-tracking features [21, 22]. Furthermore, the pupil diameter cannot be controlled consciously [10]. The pupil diameter changes with differing cognitive load—increasing with greater cognitive load and vice versa [10, 22–24]. Other external factors (e.g., light, emotions) also have an effect on the pupil diameter. Thus, controlled environments are required for an accurate measurement of cognitive load based on the pupil diameter [10]. There is already a broad range of systems available that depict on the pupil diameter for assessing real-time cognitive load, such as differentiating between truthful and deceptive answers [25], optimising time management during work [11] or optimising interruption management for delivering mails during work [8].

Cognitive load, independent from the method for measurement, has already been used in systems to manage interruptions, optimising the timing of interruptions to minimise disruption. By the usage of workload-aligned task models, right moments for interruptions can be determined [9]. [3] used smartphones mounted at each office door to display a red or green background, signalling whether the office worker is interruptible or not, by inferring cognitive load using consumer wearables. [8] identify low-workload moments by analysing the pupil diameter of users and tested their method in a single user scenario with an email-answering task with interruptions during working. Their method showed to improve performance compared to interrupting at random times.

In the next section, we provide our CLBoundaryManager concept, leveraging on the most important findings from related work.

3 A Concept for Cognitive Load Boundary Management

We present CLBoundaryManager, a boundary management concept for communication between the domains work and family. Our concept contains three components (cf. Fig. 1): a *family interface* for sending out communication requests to the family member in the work domain; a *communication gateway* for handling communication requests from the family domain to the work domain, considering the cognitive load level by the use of eye tracking of the family member in the work domain to decide on whether to pass communication requests or to delay these; and a *work interface* to accept or decline incoming communication requests.

If family members want to communicate with the working family member, they access the family interface to set up a communication request, including a relevance indication (e.g., important, neutral, not important) and optionally few describing words. They can further see the current state of the working member (e.g., being in a cognitively

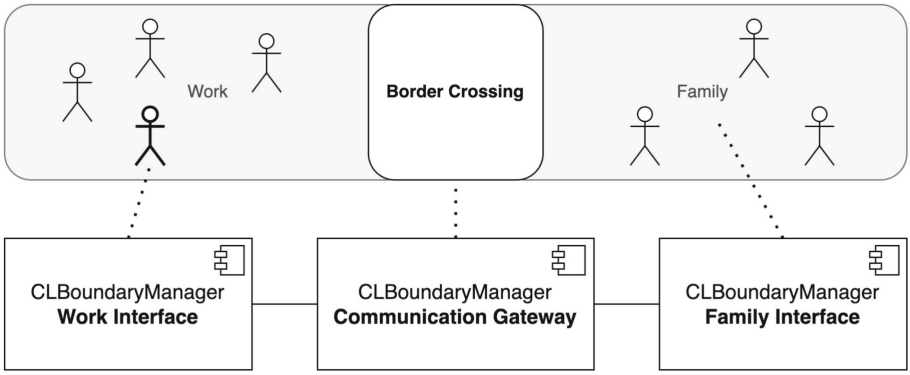


Fig. 1. The CLBoundaryManager system with its three components—the work interface, the family interface and the communication gateway.

demanding task or being potentially available interruptions), which helps them understand the appropriateness of the timing for their request. This feature draws from the concept of social cues for signalling interruptibility, adapted for a digital environment without physical proximity.

Once a communication request is submitted, the communication gateway evaluates the cognitive load of the working family member. If the cognitive load level is high and the relevance indication is low, the request will be delayed until the cognitive load level decreases. In other cases, e.g., with a lower cognitive load level or higher relevance indications, the gateway will pass the request to the work interface. This ensures that interruptions are aligned with moments of low cognitive load, minimising disruption as suggested in prior research on interruption timing optimisation. Furthermore, by targeting moments of low cognitive load, the system creates opportunities for incubation, fostering creative processes and enhancing problem-solving abilities. This approach highlights how the CLBoundaryManager not only reduces disruptions but also leverages interruptions to optimise work and thought processes.

Based on the above-described variables, the request will be displayed differently; and in any case, the working member has the chance to accept or decline the request. In case of denial, the request will be put back to the waiting queue. This approach addresses the limitations of existing 'Do Not Disturb' modes, as it incorporates context-aware interruption management and prioritisation of notifications.

The cognitive load is assessed using real-time pupil diameter measurements from eye tracking. The use of pupil diameter for assessing cognitive load is highly reliable, as it is an unconscious, real-time indicator, though its accuracy benefits from controlled environments. By interpreting the measurements as cognitive load levels, thresholds can be defined in the communication gateway to support informed decisions.

The CLBoundaryManager concept fosters better boundary management between work and family by using context-aware features to optimise the timing of interruptions, reducing family-work conflict. By leveraging these mechanisms, it minimises the impact on job satisfaction and well-being, while promoting effective communication.

Furthermore, the system aligns with the themes und constituent elements of 'presence-in-absence' by ensuring reciprocal contact, supporting expressiveness through relevance indicators, and maintaining a private, closed communication channel. The shared interface creates common ground by showing the status of the working family member (e.g., cognitive load or availability) and the presence of other family members in the system. Symbolic codes (e.g., importance levels) simplify communication, and the system remains unobtrusive by delivering requests asynchronously during low cognitive load to minimise disruptions and maintain connectedness.

In the next section we introduce the implementation of our CLBoundaryManager.

4 CLBoundaryManager Implementation

The implementation of the CLBoundaryManager system consists of two components: a client-side web application, implementing the family and work interface, and a communication gateway on server-side. Each component was developed to adhere to the conceptual design outlined earlier, focusing on context-aware communication management and seamless integration between work and family domains.

4.1 System Overview

We implemented the CLBoundaryManager system using a client-server architecture (cf. Fig. 2). The family and work interfaces are web applications developed in JavaScript Svelte (version 3.59.2) and Python (version 3.13.1), while the communication gateway operates on a Node.js (version 23.5.0) server. This architecture facilitates real-time data exchange between components, ensuring effective handling of communication requests and cognitive load assessment.

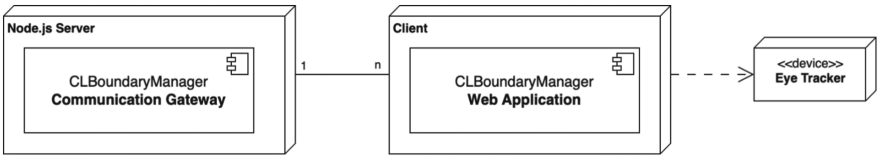


Fig. 2. System overview of CLBoundaryManager.

4.2 Family and Work Interface

While the concept introduced the family and work interface as two separate components, the implementation contains both interfaces in a single web application. When starting the web application, users can choose their role (respectively, worker or family member; cf. Fig. 3). Based on this, the user interface will be presented differently. Apart from their role, users are asked to provide their name and are required to allow notifications, so that communication requests can be properly transmitted once all requirements are fulfilled (e.g., that the cognitive load of the working family member must be below a defined threshold).

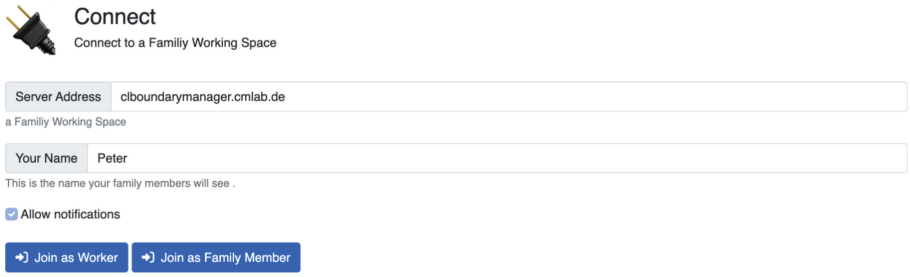


Fig. 3. The connect screen shown at startup of the CLBoundaryManager web application.

Fostering connectedness by showing all active users in the family and work domain. Both work and family interface (cf. Fig. 4) use visually distinct 'bubbles' to represent family members, fostering a sense of connectedness by providing an intuitive and personalised view of who is currently logged into the system. Each bubble is labelled with the family member's name, making it easy to identify who is available. The current user's bubble is distinctly marked to help them orient themselves within the interface. The bubbles are grouped into two life domains—work and family—clearly separated but displayed side by side, reinforcing the dual focus of the application while highlighting the current context. This dual-domain structure ensures users always understand their role and connection to others in the system.

The work interface and the family interface share a similar design language, emphasising consistency and ease of use. However, a key distinction lies in functionality: while family members can initiate communication requests through the family interface, workers cannot send such requests back. This asymmetry reflects the lack of cognitive load measurement tools (i.e., eye trackers) available to family members, which are essential for determining interruptibility. The worker's role is primarily to receive and manage requests based on cognitive load and priority. The design of the work and family interface, implemented as a responsive web application, bridges digital and social cues, ensuring accessibility and adaptability across devices while fostering a continuous sense of connectedness within the family.

Creating communication requests. Users in the family domain can create communication requests, but only for users in the work domain. A request contains a title summarising the request, a short description offering additional context, and a priority level (1–4), where 1 is the top priority and 4 is the least. Once submitted, the request is sent to the communication gateway for processing.

If the user in the work domain has a low cognitive load, the request is displayed immediately. The worker receives a push notification, and the request details are shown when opening the CLBoundaryManager web application. The worker can then either accept the request or postpone it. If postponed, the request is deferred for a fixed time (configurable in the system settings) and re-evaluated later. If accepted, the systems informs both users to meet. At the end, any of the users can end the meeting so that the user in the work domain is not blocked for further communication requests.

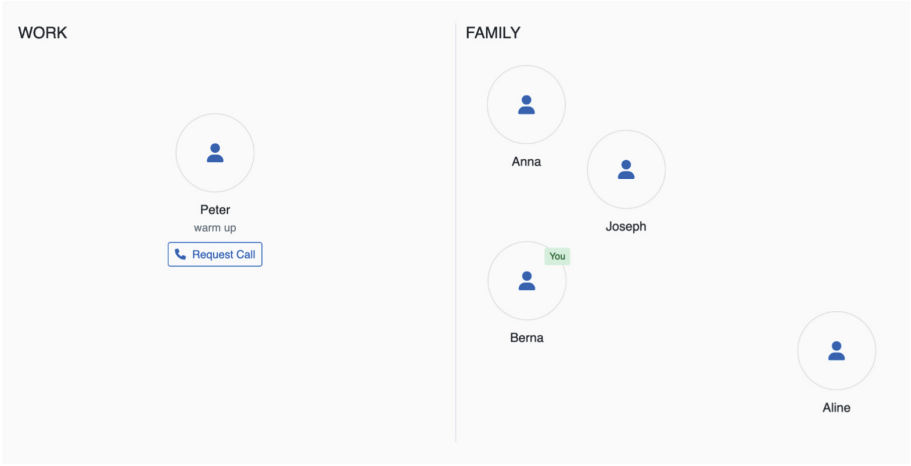


Fig. 4. Interface for user Berna in the family domain. Berna can see that Peter has just arrived in the office (warmup phase). Later Peter's bubble will show the current cognitive load.

Requests are handled differently depending on their priority. For priority 1, requests bypass cognitive load considerations and are shown immediately, regardless of the worker's current cognitive load. For remaining priorities 2–4, the requests are displayed only if the worker's cognitive load is below specific thresholds. For instance, a priority 4 request requires a lower cognitive load than priority 2. These thresholds can be configured.

The visual representation of the request (cf. Fig. 5) also varies by priority. Each priority level has a distinct colour and warning symbol, ensuring the worker can quickly gauge urgency. This differentiation is applied both in the web application and in the system's push notifications.

When multiple requests are submitted in the same timeframe (e.g., from different family members), the system ranks them by the provided priority, processing higher-priority requests first. This ensures that the most critical communication receives attention promptly, even during busy periods.

4.3 Communication Gateway Processing Communication Requests Based on Cognitive Load Data

The communication gateway is implemented as a Node.js server and is the central processing point for all communication requests. It maintains a record of all active users across both the family and work domains and contains the decision-making logic for request handling. For users in the work domain, where it is required to have an eye tracker connected, the web application transmits the user's current cognitive load level to the gateway within configurable floating time windows.

The cognitive load level is calculated based on the user's pupil diameter during the observed time window. Specifically, the current pupil diameter is compared to the minimum and maximum pupil diameters within the same window. The cognitive load

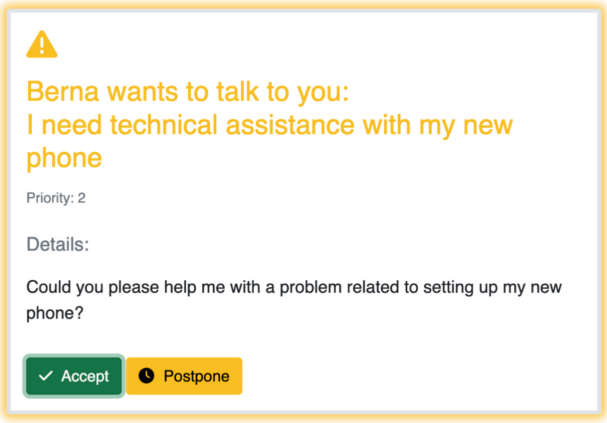


Fig. 5. A communication request displayed to a user in the work domain.

level is determined by subtracting the minimum pupil diameter from the current pupil diameter, dividing the result by the difference between the maximum and minimum pupil diameters in the time window. This results in a value between 0 and 100, with 100 representing the highest possible cognitive load.

The communication gateway is designed to be independent of the specific eye tracker hardware. While within this article, we utilise the Tobii Pro Spectrum eye tracker, the system allows for integration with other devices if the client-side implementation provides the necessary interface. This design ensures flexibility and adaptability for future use cases or hardware changes, with no impact on the gateway itself.

The gateway processes communication requests based on priority and cognitive load data. For requests with the highest priority (i.e., priority 1), the system bypasses cognitive load considerations entirely, thus delivering communication requests immediately to the worker. For the remaining priorities (2–4), cognitive load thresholds are configured for each priority level. The higher the priority, the more cognitive load a worker can sustain before the request is delayed. Requests are sorted by priority, and within the same priority, they are processed on a 'first come first serve' (FIFO) basis.

The architecture relies on real-time communication using bi-directional WebSocket channels between the clients (users in the family and work domains) and the communication gateway. Clients send data packets to the gateway containing user-specific information, the calculated cognitive load level (for work domain users), and the details of the communication request, such as title, description, and priority. The gateway processes these requests and responds to clients with relevant updates, including notifications for workers when new requests are available. This bi-directional communication ensures seamless and dynamic interaction between users and the gateway.

In the next section, we evaluate the effectiveness of CLBoundaryManager in managing interruptions between work and family domains.

5 Evaluation of the CLBoundaryManager

The introduced CLBoundaryManager concept and its implementation provides an interruption management to foster optimal communication between the two domains work and family. In this section, we will evaluate the system based on a user study to determine the effectiveness CLBoundaryManager in managing interruptions based on the cognitive load of users in the work domain.

5.1 Method

This section describes the method of the user-study, including the used materials and apparatus, an overview of the recruited participants and finally a description of the procedure with a data analysis afterwards.

We used a similar study design to [26] and [8]. [26] analysed the role of interruptions in task performance and cognitive load, which we also want to determine; and [8] evaluated a similar concept to CLBoundaryManager, as they also used the pupil diameter to postpone interruptions based on the current cognitive load of a participant.

Materials. The study included a prebriefing and informed consent. Further, the *Brief Resilience Scale* (BRS) was used to "assess the [participants'] ability to bounce back or recover from stress" [27]. We used the German version of the BRS as provided by [28]. [26] found the cognitive load to not differ in case of task interruptions for participants with medium and high resilience, because they "mitigate cognitive overload induced by task interruptions through efficient allocation of attentional resources, resilience, and effective coping strategies" [26]. By assessing the BRS in our study, we may observe possible differences on the effectiveness of CLBoundaryManager based on the individual resilience of the participants.

As main task for the participants, we used the *Continuous Performance Task* (AX-CPT). AX-APT is a broadly used paradigm in cognitive research to measure cognitive control and working memory [29, 30]. It involves sequences of five letters, where participants press "yes" only when an X probe follows an A cue (AX trials, 70%), with other letters presented in between. AY trials (10%) involve an A cue followed by a non-X probe, BX trials (10%) involve a non-A cue followed by an X probe, and BY trials (10%) involve neither an A cue nor an X probe [26, 30]. In our study we applied the AX-APT task as described by [30] and implemented it in *OpenSesame*, an "open-source, graphical experiment builder" [31].

To measure the cognitive load after each experiment (cf. Procedure), we used a German translation of the *National Aeronautics and Space Administration Task Load Index* (NASA-TLX) scale, that is capable of indirectly determining cognitive load [26, 32].

To evaluate the perceived usability of the CLBoundaryManager system, we used the German translation of the *System Usability Scale* (SUS) [33, 34], a widely adopted, technology-agnostic, and reliable tool known for its high validity and effectiveness, even with small sample sizes [35, 36].

Apparatus. We used a standard PC with Windows 10. Furthermore, a *Tobii Pro Spectrum* eye tracker (firmware version 2.6.1) was installed with 1,200 Hz. The calibration

was done with the software provided by the manufacturer (*Tobii Pro Eye Tracker Manager*, version 2.1.2) The version of *OpenSesame* for the AX-APT task was 4.0.24 and the CLBoundaryManager was run in *Microsoft Edge Browser* (version 132.0.2957.115).

Participants. We tested our implementation with 8 participants (4 female, 4 male, 0 non-binary) with age from 23 to 29 years ($M = 25.25$, $SD = 1.85$). Participants were recruited with mouth-to-mouth sampling in the university. Three participants were wearing glasses, one participant was wearing contact lenses, and no participant had former eye-surgeries with remaining scars on the cornea.

Procedure. For each participant, separate time slots were arranged. The study was conducted in a lab with identical circumstances for all participants (esp. Lighting and noise). Participants were asked to turn off their phones. The study started with the pre-briefing, followed by the informed consent and the printed BRS.

Participants then performed three experiments; accordingly to [8] we used a within-subject repeated-measurement design with three conditions in randomised order. In each experiment, participants were presented with the AX-CPT as main task. Therefore, the AX-CPT was explained once and a short practice phase (AX, AY, BX and BY each twice) was preceding each experiment.

A total of 30 trials were conducted, consisting of 70% AX trials, 10% AY trials, 10% BX trials, and 10% BY trials. Accordingly to [26], each signal was presented for 1000ms, and after each trial a fixation cross in the centre of the screen was displayed for 2000ms. A break of 90 s was implemented after every 10 trials to reduce cognitive load (in the experiment using the CLBoundaryManager system, the break facilitated the delivery of interrupting push notifications). The breaks were included across all three experiments to ensure consistent conditions. Participants were instructed to continue looking at the screen during the breaks (so measurement of the pupil diameter was possible).

In experiment I, participants did not receive any interruptions, serving as control condition. In experiment II, participants received interruptions through communication requests delivered using push messages at random times. These requests included reviewing a shopping list for later with a family member, helping set up a new laptop, planning a family celebration, and assisting in subscribing to a new newspaper. All communications had the same priority to have similar conditions for all experiments and participants. In experiment III, participants received the same interruptions; but postponed based on their cognitive load determined by the continuous measurement pupil diameter, triggered by the CLBoundaryManager system. Experiment III in contrast to the other experiments preceded with a calibration of the eye tracker and an introduction to CLBoundaryManager. After each experiment, cognitive load was measured through printed NASA-TLX. At the end of the study, each participant filled out the printed SUS.

5.2 Results and Discussion

A total of eight experimental data were collected in this study, of which descriptive statistics of cognitive load, task performance and reaction time are shown in Table 1.

We used a repeated measures ANOVA, but did not find any statistically significant differences across the experimental conditions for any of the dependent variables. For

Table 1. Descriptive statistics for the three experiments.

Experiment	NASA-TLX score (M, SD)	Mean RT ms (M, SD)	Mean ACC % (M, SD)
Experiment I (no interruptions)	6.46, 3.40	888.55, 170.18	97.76, 3.16
Experiment II (random interruptions)	6.25, 3.13	840.47, 140.04	96.27, 2.80
Experiment III (IMS interruptions)	6.52, 2.63	947.55, 332.10	97.56, 1.62

mean accuracy, no significant effect of interruption type was observed ($F(2, 14) = 0.611, p = 0.556, \eta^2 = 0.060$). For mean reaction time, the analysis showed no significant differences across conditions ($F(2, 14) = 0.333, p = 0.723, \eta^2 = 0.035$). For the measured workload after each experiment (i.e., NASA-TLX), there was no significant variation between conditions ($F(2, 14) = 0.074$, $p = 0.929$, $\eta^2 = 0.001$). In all cases, the effect sizes were small to negligible, indicating that the experimental manipulations explained only a minimal proportion of the variance in the dependent variables. The assumptions of sphericity were met for all measures ($\epsilon > 0.8$).

These findings indicate that the type of interruptions, whether randomly timed or managed by CLBoundaryManager as interruption management system (IMS), did not significantly influence accuracy, reaction time, or perceived workload in this context. However, several considerations must be addressed when interpreting these results.

The simplicity and brevity of the tasks used in this study may have limited the potential for interruptions to disrupt performance or increase workload.

Future studies should investigate these effects in more complex or dynamic task environments, such as those involving multi-tasking or time-critical decision-making.

Furthermore, while IMS did not demonstrate clear benefits over random interruptions in this study, it is possible that its advantages may manifest more strongly in high-stakes domains or real-world applications. Another problem might have been, that in experiment III some of the participants did also receive interruptions during the AX-CPT task and not during breaks (as it was intended).

In the future, we will try to determine possible reasons for the current concept working better for some participants than for others, and based on these findings improve CLBoundaryManager. Reasons might also be found in the study design, since participants reported that they thought about non-study related things during the breaks (i.e., mind wandering), which might also have caused higher cognitive load during breaks and thus leading to non-delivery of communication requests during these.

In future studies, it may be further interesting to see whether the resilience of a participant correlates with the effectiveness of interruption management systems. Studies show that probands with higher resilience scale perceive less stress and have a higher well-being, also in workspaces [37]. The participants of our study all had normal or high resilience (6 normal, 2 high; $M = 3.93$, $SD = 0.44$). A study with participants with a

low resilience might hold enriching insights to improve the current system logic. Due to the low sample size of our study, we did not further separate our measured data.

We reached a 'good' ([36, 38]) SUS score for the CLBoundaryManager ($M = 75.94$, $SD = 5.86$), with a grade of grade of B (according to [36, 39]). Most users perceived the system as not too complex, easy to use, and felt able to use the system without additional help by experts. In the future, the integration of key functions should be considered for optimisation ('I found the various functions in this system were well integrated' with $M = 2.75$, $SD = 0.43$, whereby zero would be the worst value and four the best achievable). Also, the design of the system might be modified to foster a higher confidentiality with the usage ('I felt very confident using the system' with $M = 2.38$, $SD = 0.70$), e.g. through improving the onboarding process. Mixed feedback with a high standard deviation was provided for the question, whether users like to 'would like to use this system frequently' ($M = 2.13$, $SD = 1.27$). Since we conducted the study in a lab with a non-realistic task scenario, it would be interesting to see how this value differs within a study in the field (i.e., using CLBoundaryManager over a longer time during real work in the field). Extending the findings of this study with future findings may reveal interesting insights in which groups of users see potential in such an interruption management system optimised for the communication between work and family domain. It can also be interesting to compare the acceptance of such a system in other domains, e.g., work-work domain.

6 Conclusions and Future Work

We introduced the CLBoundaryManager as a novel system to manage interruptions between work and family domains, leveraging real-time cognitive load assessments through eye-tracking technology. By aligning the timing of interruptions with moments of lower cognitive load, this approach minimises disruption during demanding tasks while fostering connectedness between domains. The system's design, including context-aware decision-making and a dual-interface setup, demonstrated its potential to balance productivity and social interaction. Initial findings indicate that the system is well-received, with users highlighting its usability and capacity to reduce unnecessary interruptions. However, limitations such as the need for controlled environments and minimal observed effects in the user study suggest room for improvement.

Future work will focus on addressing these limitations through a more comprehensive study with a larger sample size and a more long-term task design, potentially including real-world scenarios to better reflect actual user behaviour. Another key development is the integration of urgency detection derived directly from text inputs, eliminating the need for users to manually assign priorities. Furthermore, additional factors influencing interruptibility will be explored, such as linking the system to user calendars in the work domain to account for scheduled commitments. These enhancements aim to increase the system's robustness, usability, and versatility, ensuring its broader applicability across different domains and contexts.

Acknowledgments. We thank the members of the Cooperative Media Lab at the University of Bamberg. We also thank the anonymous reviewers for insightful comments.

References

1. Gross, T., von Kalben, M.: A Literature Review on Positive and Negative Effects of Interruptions and Implications for Design. In: Proceedings of the IFIP Conference on Human-Computer Interaction. Springer (2023). https://doi.org/10.1007/978-3-031-42293-5_38
2. Gillie, T., Broadbent, D.: What makes interruptions disruptive? A study of length, similarity, and complexity. Psychol. Res. **50**(4), 243–250 (1989). https://doi.org/10.1007/BF00309260
3. Schaule, F., Johanssen, J.O., Bruegge, B., Loftness, V.: Employing consumer wearables to detect office workers' cognitive load for interruption management. Proc. ACM. Interact. Mob. Wear. Ubiquit. Technol. **2**(1), 1–20 (2018). https://doi.org/10.1145/3191764
4. Wei, R., Lo, V.-H.: Staying connected while on the move: Cell phone use and social connectedness. New media & society **8**(1), 53–72 (2006). https://doi.org/10.1177/1461444806059870
5. McDaniel, B.T., Drouin, M.: Daily technology interruptions and emotional and relational well-being. Comput. Hum. Behav. **99**(2019), 1–8 (2019). https://doi.org/10.1016/j.chb.2019.04.027
6. Steele, J.: What is (perception of) effort? Objective and subjective effort during task performance. *PsyArXiv* (2020). https://doi.org/10.31234/osf.io/kbyhm
7. Goyal, N., Fussell, S.R.: Intelligent interruption management using electro dermal activity based physiological sensor for collaborative sensemaking. Proc. ACM interact. Mob. Wear. Ubiqui. Technol. **1**(3), 1–21 (2017). https://doi.org/10.1145/3130917
8. Katidioti, I., et al.: Interrupted by your pupil: An interruption management system based on pupil dilation. Int. J. Hum. Comput. Interact. **32**(10), 791–801 (2016). https://doi.org/10.1080/10447318.2016.1198525
9. Iqbal, S.T., Bailey, B.P.: Investigating the effectiveness of mental workload as a predictor of opportune moments for interruption. In: Proceedings of the CHI'05 extended abstracts on Human factors in computing systems. (2005). https://doi.org/10.1145/1056808.1056948
10. Eckstein, M.K., Guerra-Carrillo, B., Miller, A.T., Bunge, S.A.: Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development? Dev. Cogn. Neurosci. **25**(2017), 69–91 (2017). https://doi.org/10.1016/j.dcn.2016.11.001
11. Maleck, M., Gross, T.: CoLoTiMa: A Cognitive-Load Based Time Management Tool. Proc. Mensch und Comput. **2024**, 1 (2024). https://doi.org/10.1145/3670653.3677485
12. Gregory Trafton, J., Monk, C.A.: Task interruptions. Rev. hum. Fact. Ergon. **3**(1), 111–126 (2007). https://doi.org/10.1518/155723408X299852
13. Gross, T., Mueller, A.: NotificationManager: Personal boundary management on mobile devices. In: Proceedings of the Human-Computer Interaction–INTERACT 2021: 18th IFIP TC 13 International Conference, Bari, Italy, August 30–September 3, 2021, Proceedings, Part IV 18. Springer (2021). https://doi.org/10.1007/978-3-030-85610-6_15
14. Jarupreechachan, W., Kitchat, K., Surasak. T.: 'Do Not Disturb': An Implication Design To Be Alerted But Less Stress. In: Proceedings of the 2023 International Seminar on Application for Technology of Information and Communication (iSemantic). IEEE (2023). https://doi.org/10.1109/iSemantic59612.2023.10295289
15. Qin, Y., Bhattacharya, T., Kulik, L., Bailey, J.: A context-aware do-not-disturb service for mobile devices. In: Proceedings of the 13th International Conference on Mobile and Ubiquitous Multimedia (2014). https://doi.org/10.1145/2677972.2678003
16. Leroy, S., Schmidt, A.M., Madjar, N.: Working from home during COVID-19: A study of the interruption landscape. J. Appl. Psychol. **106**(10), 1448 (2021). https://doi.org/10.1037/apl0000972

17. Höge, T.: When work strain transcends psychological boundaries: An inquiry into the relationship between time pressure, irritation, work–family conflict and psychosomatic complaints Stress and Health. J. Int. Soc. Invest. Stress **25**(1), 41–51 (2009). https://doi.org/10.1002/smi.1226
18. Netemeyer, R.G., Boles, J.S., McMurrian, R.: Development and validation of work–family conflict and family–work conflict scales. J. Appl. Psychol. **81**(4), 400 (1996). https://doi.org/10.1037/0021-9010.81.4.400
19. Allen, T.D., Cho, E., Meier, L.L.: Work–family boundary dynamics. Annu. Rev. Organ. Psychol. Organ. Behav. **1**(1), 99–121 (2014). https://doi.org/10.1146/annurev-orgpsych-031413-091330
20. Howard, S., Kjeldskov, J., Skov, M.B., Garnæs, K., Grünberger, O.: Negotiating presence-in-absence: contact, content and context. Proc. SIGCHI Conf. Hum. Fact. Comput. Syst. (2006). https://doi.org/10.1145/1124772.1124906
21. Zagermann, J., Pfeil, U., Reiterer, H.: Measuring Cognitive Load using Eye Tracking Technology in Visual Computing. Proceedings of (2016). ACM Press. (2016). https://doi.org/10.1145/2993901.2993908
22. Zagermann, J., Pfeil, U., Reiterer. H.: Studying eye movements as a basis for measuring cognitive load. In: Proceedings of the Extended Abstracts of the 2018 CHI conference on human factors in computing systems (Montreal QC Canada) (2018). https://doi.org/10.1145/3170427.3188628
23. Pfleging, B., Fekety, D.K., Schmidt, A., Kun, A.L.: A model relating pupil diameter to mental workload and lighting conditions. In: Proceedings of the CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (San Jose California, USA) (2016). https://doi.org/10.1145/2858036.2858117
24. Dionisio, D.P., Granholm, E., Hillix, W.A., Perrine, W.F.: Differentiation of deception using pupillary responses as an index of cognitive processing. Psychophysiol. **38**(2), 205–211 (2001). https://doi.org/10.1111/1469-8986.3820205
25. Maleck, M., Gross, T.: AnswerTruthDetector: A Combined Cognitive Load Approach for Separating Truthful from Deceptive Answers in Computer-Administered Questionnaires. i-com - Journal of Interactive Media **22**(3), 241–251 (2023). https://doi.org/10.1515/icom-2023-0023
26. Yuan, X., Zhong, L.: Effects of multitasking and task interruptions on task performance and cognitive load: considering the moderating role of individual resilience. Curr. Psychol. **2024**, 1–11 (2024). https://doi.org/10.1007/s12144-024-06094-2
27. Smith, B.W., Dalen, J., Wiggins, K., Tooley, E., Christopher, P., Bernard, J.: The brief resilience scale: assessing the ability to bounce back. Int. J. Behav. Med. **15**(2008), 194–200 (2008). https://doi.org/10.1080/10705500802222972
28. Chmitorz, A., Wenzel, M., Stieglitz, R.-D., Kunzler, A., Bagusat, C., Helmreich, I., Gerlicher, Anna, Kampa, M., Kubiak, T., Kalisch, R.: Population-based validation of a German version of the Brief Resilience Scale. PloS one **13**(2), e0192761 (2018). https://doi.org/10.1371/journal.pone.0192761
29. Braver, T.S., Barch, D.M.: A theory of cognitive control, aging cognition, and neuromodulation. Neurosci. Biobehav. Rev. **26**(7), 809–817 (2002). https://doi.org/10.1016/S0149-7634(02)00067-2
30. Berry, G.M.: Processing linguistic variation through dual mechanisms of cognitive control. Linguist. Vanguard **2**(s1), 20160033 (2016). https://doi.org/10.1515/lingvan-2016-0033
31. Mathôt, S., Schreij, D., Theeuwes, J.: OpenSesame: An open-source, graphical experiment builder for the social sciences. Behav. Res. Methods **44**(2012), 314–324 (2012). https://doi.org/10.3758/s13428-011-0168-7

32. Hart., S.G.: NASA-task load index (NASA-TLX); 20 years later. In: Proceedings of the Proceedings of the human factors and ergonomics society annual meeting. Sage publications Sage CA, Los Angeles, CA (2006). https://doi.org/10.1177/154193120605000909
33. Brooke, J.: SUS: A quick and dirty usability scale. CRC Press, London. (1996). https://doi.org/10.1201/9781498710411
34. Rummel, B., Ruegenhagen, E., Reinhardt, W.: System usability scale (translated into german) (2014)
35. Vlachogianni, P., Tselios, N.: Perceived usability evaluation of educational technology using the System Usability Scale (SUS): A systematic review. J. Res. Technol. Educ. **54**(3), 392–409 (2022). https://doi.org/10.1080/15391523.2020.1867938
36. Brooke, J.: SUS: a retrospective. J. Usability studies, **8**(2) (2013). https://doi.org/10.5555/2817912.2817913
37. Kermott, C.A., Johnson, R.E., Sood, R., et al.: Is higher resilience predictive of lower stress and better mental health among corporate executives? PloS one **14**(6), e0218092 (2019). https://doi.org/10.1146/annurev-orgpsych-031413-091330
38. Bangor, A., Kortum, P., Miller, J.: Determining what individual SUS scores mean: Adding an adjective rating scale. J. Usabil. Stud. **4**(3), 114–123 (2009). https://doi.org/10.5555/2835587.2835589
39. Sauro, J.: A practical guide to the system usability scale: Background, benchmarks & best practices. Measuring Usability LLC (2011)

Chinese Talk About Depression on Tiktok and Its Deliberative Differences from the Russian Online Dialogue

Alena Rodicheva and Svetlana S. Bodrunova(✉)

St. Petersburg State University, 7–9 Universitetskaya nab., St. Petersburg 199004, Russia
s.bodrunova@spbu.ru

Abstract. In this pilot study we look into how the discussion of depression in China stands from the perspective of modern deliberative theory. We build upon our previous research on brand cumulative crises where we have demonstrated the cumulative nature of online crises experienced by two Russian brands and studied the opinion patterns inside the discussion around one of the brands. For the current paper, we study the discussion of depression under one of the videos in Chinese TikTok - there are many such discussions, they need to be studied in an automated way, but first we need to look at how exactly Chinese citizens discuss depression, what are the features. 227 comments were coded manually with Kappa testing (agreement level 0.7 to 0.9). Our findings show that most of the comments of the Chinese about depression are cumulative, directed to the post's author or *urbi et orbi*, there is no discussion at all. No comments are directed to authorities and there is a tiny percentage of users' interaction: sharing advice or expression of sympathy. There is a vacuum of agreement and disagreement both with interlocutors and authorities, no deliberative pattern, only cumulative one. Besides, almost all discussion participants define depression as a serious illness and mostly rely on personal experience. There are no reliable sources and almost no scientific data mentioned which does not make these discussions a trustworthy source of medical information. There is no irony/sarcasm and aggression in the comments which creates a safe space for sharing personal stories and being vulnerable with strangers. Chinese public dialogue seems to be very different from the Russian one in terms of deliberative patterns. Russian crisis communication case is mostly cumulative, however there are some polarized micro-dialogues (chains of <10 comments), more comments are directed to interlocutors; higher degree of disagreement. The study adds to today's understanding of the nature of Chinese public dialogue about depression and might be useful for the patients' treatment.

Keywords: depression · public dialogue · opinion patterns · cumulative deliberation · social media · online discussion

1 Introduction

Given the insufficient and uneven distribution of mental health resources and the intense stigma attached, an increasing number of Chinese patients with mental disorders are turning to online support groups (OSGs) for advice [1]. Many of them are also engaging

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 351–359, 2025.
https://doi.org/10.1007/978-3-031-93539-8_23

in discussions while commenting under depression-related content on social media, however the traditional way of discussing these issues might differ from European, American and Russian practices.

Habermas's deliberative democracy theory says that conversation should be rational, dialogic, and aimed at consensus [2]. However, a newer theory suggests that affective publics [3] within the framework of cumulative deliberation [4] can be emotional, aggressive, irrational, non-dialogical, poorly grounded in argument, conflictual, polarized, opinionated, fragmented and, most important, not seeking any meaningful consensus or even temporary agreement. We also know that, in Russia and other countries, complaints addressed to authorities on various issues are often voiced on social media [5]. In this pilot study, we look at whether health issues evoke deliberation and policy discussion in cultures with non-Western traditions of public dialogue. Thus, we pose the following *research questions*:

RQ1. Is the nature of the discussion about depression more deliberative or more cumulative, and are there any patterns of the discussion development?

RQ2. How do Chinese people define depression and talk about it?

RQ3. How does Chinese discussion differ from the Russian online public dialogue?

To address the above listed RQs, we build upon our previous research on brand cumulative crises. In it, we have proven the cumulative nature of online crises experienced by two Russian brands [6] and studied the opinion patterns inside the discussion around one of them [7, 8].

For the current paper, we study the discussion of depression under one of the videos in Chinese TikTok - there are many such discussions, they need to be studied in an automated way, but first we need to look at how exactly Chinese citizens discuss depression, what are the features. 227 comments were coded manually with Kappa testing (agreement level 0.7 to 0.9) [9]. The questions in the coding sheet are the same ones we used when studying online discussions from a deliberative perspective, as well as when studying cumulative brand crises:

1. Deliberative features: a) addressee of the appeal, b) expression of (dis)agreement with the interlocutor, c) expression of (dis)agreement with the government policy;
2. Questions about depression: a) description and perception of depression, b) presence of personal experience, c) presence of experience of relatives and friends;
3. Characteristics of discourse: a) presence of a reliable source, b) presence of scientific data, c) presence of irony/sarcasm, d) presence of aggression.

The paper is organized as follows. In Sect. 2, we provide theory about cumulative deliberation and the features of online discussions. In Sect. 3, we describe our methods. Section 4 presents our results. In Sect. 5, we make conclusions.

2 The Conceptual Framework of Cumulative Deliberation vs. Discussing Depression on Chinese Social Media

2.1 The Cumulative Nature of Online Discussions

The quality of public discussions, according to some researchers, tend to decrease because of public spheres hybridization [10]. The growth in the number of the user-generated content is also seen as a major factor. The deliberation theory by Juergen Habermas [2] suggests that participants of an online discussion seek to reach a meaningful consensus. However, we believe that this is not true in most of the cases of modern public dialogues online, the current theory of the public sphere imposes very high and unrealistic expectations on participants in political discussions, user speech may be meaningless, biased, emotional, poorly grounded in argument and not seeking any consensus or even temporary agreement.

The conceptual framework of cumulative deliberation [4] describes socially-mediated speech as 'constantly dynamic' and mostly non-deliberative – antagonistic [11], irrational, emotional, conflictual, polarized, and non-consensual.

We believe that this concept that is closely related to such system theories as the he theory of dissipative systems [12] and the affective publics [3], might be used for analyzing various online discussions. In this study, we will try to apply the concept of cumulative deliberation to study the Chinese public dialogue about depression.

2.2 Discussions on Depression on Chinese Social Media

Millions of people around the world suffer from depression nowadays. In China, this issue is extremely important. Chinese researchers emphasize the role of Chinese cultural beliefs in the power of inner self-control and traditional medicine in patients' understandings of depression, illness management, and social interactions. It might cause problems of understanding depression in society and make patients marginalized, suffer discrimination, and lose their jobs [13].

The dilemma of disclosing one's illness experiences in real life has led to a proliferation of online health communities [14]. One of the studies explores the evolutionary pattern of discussion topics over time in a Chinese online depression self-help community. Authors identify the discussion topics of the group that includes self-reflection, daily record, peer diagnosis, companionship support and instrumental support [15]. The effectiveness of such communities has also been proved by Chinese researchers who think that the internet-based self-help intervention can be effective in treating depression in adolescents and young adults [16].

Even though there are many studies on that topic, none of them look into the structure of online discussion about depression which can be done with the help of the conceptual framework of cumulative deliberation. Moreover, comparing the Chinese public dialogue with the study of a Russian cumulative crisis case might help to see to what extent the depression related talk in China is conflictual and whether or not it can turn into a more political conversation through directing complaints about poor mental health policy to the authorities.

3 Method

We build upon our previous research on brand cumulative crises [7]. There, we have demonstrated the cumulative nature of online crises experienced by two Russian brands and studied the opinion patterns inside the discussion around one of the brands. The datasets for the research were collected both manually and automatically. For the current paper, we study the discussion on depression under one video on Chinese TikTok. There are many such discussions, and they need to be studied in an automated way, but first we need to look at how the Chinese context may shape the talk, to integrate this knowledge into automated research design. 227 comments that contained the word 'depression' were coded manually with Kappa testing (intercoder agreement 0.7 to 0.9).

The coding variables are similar to those we used when studying online discussions within cumulative brand crises in Russia from a deliberative perspective (see below). Among them, there are the deliberative features: 1) addressee of the appeal, 2) expression of agreement/disagreement with the interlocutor, 3) expression of agreement/disagreement with the government policy; and questions on depression: 1) description and perception of depression, 2) presence of personal experience, 3) presence of experience of relatives and friends; characteristics of discourse: 1) presence of a reliable source, 2) presence of scientific data, 3) presence of irony/sarcasm, 4) presence of aggression.

Then, we visualised collected and coded data by using the interpretative reading method and building graphs presented in the next chapter and made conclusions that help us answer our research questions.

4 Results

RQ1. To find out if cumulative or deliberative patterns are dominant in the discussion, we manually coded the addressee of the appeal section and built a graph (Fig. 1) showing how many comments were directed to the post's author (1), another commentator (2), authorities (3), urbi et orbi (4). We also coded the comments that had no clear direction towards anyone (99).

As we can see, most of the comments of the Chinese about depression appear to be cumulative, there is no deliberative pattern at all. Almost all comments are directed to the post's author or urbi et orbi, there is no discussion at all. Besides, there are completely no comments directed to the authorities which shows that the Chinese do not tend to blame the government for creating the environment that can cause citizens' depression. We also noticed vacuum of agreement and disagreement both with interlocutors and authorities. There is only a tiny percentage of users' interaction: sometimes people are sharing advice or expressing sympathy towards someone's story, but comments like that are hard to find.

However, there are some examples of users' interaction, such as:

Comment 1: *I was under too much pressure and I couldn't hold back any longer. My boss was the first to notice that there was something wrong with my mental state. She also asked my colleagues if they had noticed anything wrong with my recent emotional state. At that time, I also asked for leave to undergo a mental health check. The doctor*

diagnosed me with moderate depression. The doctor also said that if you control your emotions and take medication, you will gradually get better.

Comment 2: *Insomnia can also cause depression and anxiety. Now I can't say whether it's depression from insomnia or insomnia from depression. The doctor also prescribed sleeping pills; I can't sleep without them.*

Comment 3: *I'm about the same age as you and I've experienced exactly the same symptoms.*

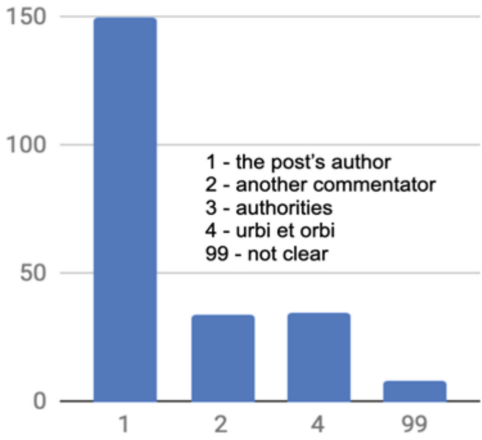


Fig. 1. Addressee of the appeal

RQ2. To answer the research question about the definition of depression by the Chinese, we looked closely into the text of comments and found several definitions using the interpretative reading method. As you can see in the graph (Fig. 2), there were examples of the word 'depression' being used but not described (0). Besides, it was described as: disease (1), social problem (2), social trait (3), laziness or an individual character trait (4). There were also some combined descriptions (99) and other conceptualisations (999), but they were very few and could not shape our understanding of how the Chinese define depression.

As one can clearly see, almost all commentators define depression is a serious illness that needs to be treated, other conceptualisations were not used significantly. Here are some examples of how users talk about depression:

Comment 1: *It's terrible to feel no emotion. Some people around me think I'm sad, but in fact I'm not happy or sad, I can't perceive any emotion, what you're interested in, now I don't feel any pleasure from communication, even if his face is a familiar smile, the brain does not feel any excitement at all.*

Comment 2: *Due to the lack of dopamine and serotonin in patients with depression, the brain blocks the reward function. In other words, when they encounter something, they will not give it a colorful meaning like normal people.*

Comment 3: *Traditional Chinese medicine believes that depression is the result of problems with the body. Modern medicine associates it with environmental factors, stress, as well as hormonal imbalance, protein secretion, and the immune system.*

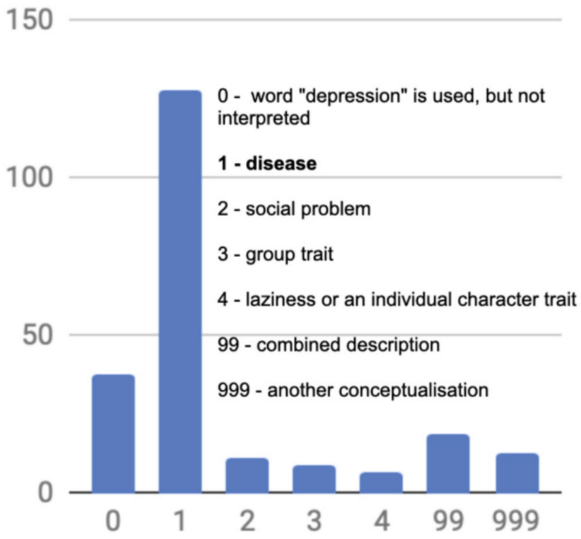


Fig. 2. Interpretations of depression.

Most of the comments are quite long as people are sharing personal stories and information they know about depression. In that rare cases when people interact with each other, they encourage their interlocutors to consult a doctor in case of depression symptoms and keep taking medication. We also found out that commentators mostly rely on personal experience (54%), sometimes their relatives and friends experience 6%). Even though there is no discussion, learning about other people's experience with depression might help some users feel they are not alone with this problem and encourage them not to give up fighting it.

Neither have we discovered any reliable sources that users link to (0%); there was almost no scientific data mentioned by the commentators (6%). This evidently lowers the deliberative potential of the discussion, as it is not based on trustworthy sources and, thus, cannot be itself a trustworthy source of information. Here are a few examples of what can possibly be considered as scientific data:

Comment 1: *Let's put it this way, it is estimated that 8% of teenagers suffer from depression, and on the Internet, this figure is even higher than 15%. Basically, for every ten Internet users I meet, I feel like there are five or six people with depression. And those I know are moderate. The above mentioned two or three Internet users are depressed.*

Comment 2: *The average age of patients with depression is definitely under 25. Most of them are teenagers, but they are not sure whether they are really sick. In their 20s, they can already go to the doctor themselves.*

Comment 3: *Although depression and anxiety are both mental health problems, they differ in terms of emotional experience, behavioral indicators, cognitive model, etc. The main characteristics of depression are persistent low mood and loss of interest in life, while the main characteristics of anxiety disorders are persistent worry and fear. In the somatic form, depression is usually characterized by hand tremors, loss of appetite, and body pain. A large amount of internal weight gain or loss may occur in a short period of time, while anxiety disorder is a symptom of autonomic nervous system disorder, such as palpitations, sweating, shortness of breath, etc.*

We also need to say that have not seen any irony/sarcasm and aggression in the comments which creates a safe space for sharing personal stories and being vulnerable with strangers.

RQ3. To give the answer to the question regarding the similarities and differences between the chinese online talk about depression and the russian public dialogue related to cumulative crises, we qualitatively juxtaposed the results of this study with our 2024 study on opinion pattern in cumulative crises of brands on social media [7]. Despite they have different topicality, which makes them less comparable than one would wish, but they also have much in common, which makes a qualitative look quite useful for designing automated quantitative research in future. The deliberative/cumulative features of the Russian and Chinese discourse may be summarized in the following way (Table 1):

Table 1. Deliberative characteristics of issue-oriented online talk in China and Russia.

Deliberative feature	China	Russia
Deliberation pattern	Nearly purely cumulative	Cumulative with micro-dialogues (chains of < 10 comments)
Addressee	Interlocutor (15%)	Interlocutor (53%),
	Author of the post (66%)	Author of the post (29%)
	Urbi et orbi (15%)	*Urbi et orbi* (17%)
Agreement/disagreement	Vacuum of both agreement and disagreement with both interlocutors and authorities	Disagreement localized in micro-dialogues; 31% of comments contain disagreement
Aggression	0%	2%*

Note. * - The low number of aggressive comments may be explained by potential cleaning out of comments by the account holder

As one can see, the Chinese case appears to be almost purely cumulative, unlike the Russian one where we saw micro-deliberative patterns. In Russia, more comments were directed to interlocutors, thus the discussion bore a more deliberative character – however, with a much higher degree of disagreement.

5 Discussion and Conclusion

Analyzing the Chinese public dialogue about depression on TikTok, we have come to the following conclusions.

RQ1. Most of the comments of the Chinese about depression are cumulative, directed to the post's author or *urbi et orbi*, there is no discussion at all. No comments are directed to authorities and there is a tiny percentage of users' interaction: sharing advice or expression of sympathy. There is a vacuum of agreement and disagreement both with interlocutors and authorities, no deliberative pattern, only cumulative one.

RQ2. Almost everyone thinks depression is a serious illness that needs to be treated. Commentators mostly rely on personal experience, sometimes their relatives and friends experience. Learning about other users' problems might help feel better. There are no reliable sources and almost no scientific data mentioned which does not make these discussions a trustworthy source of medical information. There is no irony/sarcasm and aggression in the comments which creates a safe space for sharing personal stories and being vulnerable with strangers.

RQ3. Chinese public dialogue seems to be very different from the Russian one in terms of deliberative patterns. Russian crisis communication case is mostly cumulative, however there are some polarized micro-dialogues (chains of < 10 comments), more comments are directed to interlocutors, with higher degree of disagreement. The Chinese commenting culture looks very different, it demonstrates avoidance of confrontation, but not via deliberatively reached agreement but via avoiding dialogue whatsoever. It also virtually abstains from scientific-based arguments, and 'discussion' actually turns into massive sharing of experience of depression and its overcoming, without policy discussion or criticism. However, we need to analyze a more relevant Russian case to make more precise conclusions.

The study adds to today's understanding of the nature of Chinese public dialogue about depression and might be useful for the patients' treatment.

Acknowledgements. This work has been supported by Center for International Media Research of St. Petersburg State University, project JSF HUST 121071140 (2024).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Pun, J., Joyce Yu, Q.: Sharing peer advice: a case study of Chinese patients' narratives in an online support group for mental health. Lang. Health **1**(2), 79–88 (2023). ISSN 2949–9038
2. Habermas, J.: Political communication in media society: does democracy still enjoy an epistemic dimension? the impact of normative theory on empirical research. Commun. Theory **16**(4), 411–426 (2006)
3. Papacharissi, Z.: Affective Publics: Sentiment, Technology, and Politics. Oxford University Press (2015)

4. Bodrunova, S.S.: The concept of cumulative deliberation: linking systemic approaches to healthier normativity in assessing opinion formation in online discussions. J. Assoc. Inf. Sci. Technol. 1–14 (2023)
5. Nigmatullina, K., Bodrunova, S.S., Rodossky, N., Nepiyushchikh, D.: Discourse of complaining on social networks in Russia: cumulative opinions vs. decentering of institutions. In: Antonyuk, A., Basov, N. (eds.) Networks in the Global World VI. NetGloW 2022. Lecture Notes in Networks and Systems, vol. 663, pp. 3–20. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-29408-2_1
6. Rodicheva, A., Bodrunova, S.S., Blekanov, I.S., Tarasov, N., Belyakova, N.: Crisis communication and reputation management of russian brands on social media. In: Bakaev, M., Bolgov, R., Chugunov, A.V., Pereira, R., R, E., Zhang, W. (eds.) Digital Geography. IMS 2023. Springer Geography. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-67762-5_36
7. Rodicheva, A., Bodrunova, S.S., Blekanov, I.S., Tarasov, N.: Opinion patterns in cumulative crises of brands on social media. In: Coman, A., Vasilache, S. (eds.) Social Computing and Social Media. HCII 2024. Lecture Notes in Computer Science, vol. 14704. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-61305-0_23
8. Bodrunova, S., Rodicheva, A., Blekanov, I., Ashraf, A., Yasin, Z.: Inside cumulative crises: types of consensus and users' deliberative roles in discussion outbursts around brands. In: ICTIS 2025 Proceedings, Bangkok, Springer LNNS (2025). ISSN: 2367–3370
9. Krippendorff, K.: Content Analysis: An Introduction to its Methodology. Sage Publications, London (2018)
10. Chadwick, A.: The Hybrid Media System: Politics and Power. Oxford University Press (2013)
11. Mouffe, C.: The Democratic Paradox. Verso, London (2000)
12. Prigogine, I.: Dissipative structures, dynamics and entropy. Int. J. Quantum Chem. **9**(S9), 443–456 (1975)
13. Li, G., Zhou, X., Lu, T., Yang, J., Gu, N.: Understanding depression in a chinese online community. In: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16). Association for Computing Machinery, New York, NY, USA, 515–526 (2016)
14. Li, X., Xu, K.: Enacting an online depression community in China: the community of practice perspective. Health Commun. 1–10 (2025)
15. Sun, H., Fichman, P.: Evolution of discussion topics on an online depression self-help group. Libr. Hi Tech **42**(4), 1347–1365 (2024)
16. Ma, Q., Shi, Y., Zhao, W., et al.: Effectiveness of internet-based self-help interventions for depression in adolescents and young adults: a systematic review and meta-analysis. BMC Psychiatry **24**, 604 (2024)

Are Strength-Related Episodes Provided by Others Useful for Improving Positive Emotions? Toward the Design of a Computer-Based Self-Esteem Enhancement System

Tomoro Watababe(✉), Masayuki Ando, Kouyou Otsu, and Tomoko Izumi

Ritsumeikan University, 2-150 Iwakura-cho, Ibaraki, Osaka, Japan
is0644vi@ed.ritsumei.ac.jp, {mandou,k-otsu,
izumi-t}@fc.ritsumei.ac.jp

Abstract. Mental health problems are a global social issue. Prior research has shown that young people, in particular, have low self-esteem and self-affirmation, and many support proposals have incorporated Information and Communication Technology. However, most of those proposals focus on self-care. On the other hand, in the field of psychology, recall of "episodes" (memories of personal experiences) is one factor that can increase self-esteem. It has also been shown that being praised for one's strengths by others can lead to learning new aspects of oneself. In this study, we propose a framework that allows users to exchange messages with others that focus on the partner's strengths and episodes related to those strengths. We conducted an experiment to consider two unresolved topics related to realizing the proposed framework, the specific functions required for the framework and the impact of the episodes given by others on self-esteem and mood. As a result of the experiment, we found that the task of praising each other had a positive effect on self-esteem and mood. However, no significant effect was confirmed on the effects of messages containing episodes. We found that being praised from multiple perspectives in the content of the message had a positive effect on self-esteem and mood. These results provide insights into designing a new system to support self-esteem based on praise exchange.

Keywords: Mental Health · Well-Being · Text Communication · Episode · Self-Esteem

1 Introduction

1.1 Research Background

Maintaining mental health is a major challenge worldwide. Young people tend to be influenced by social comparisons with others, both in real life and on social media, which can lead to low self-esteem [1], and low self-esteem has a strong impact on mental illness [2]. Therefore, supporting enhanced self-esteem in young people could help reduce the risk of mental illness [3].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 360–374, 2025.
https://doi.org/10.1007/978-3-031-93539-8_24

In this study, we explore the use of technological support to enhance self-esteem through feedback from others with whom people interact in their daily lives. Mobile applications designed for mental health interventions have become increasingly widespread [4, 5] and in the field of human-computer interaction (HCI), the effective design for such computer-based mental healthcare has been discussed [6, 7]. Existing computer-based interventions aimed at enhancing self-esteem often utilize cognitive-behavioral therapy (CBT) techniques to identify and modify self-critical thoughts [4, 5, 8]. Other studies have incorporated positive psychology approaches, such as keeping a gratitude journal, to emphasize the positive aspects of individuals' daily lives [9]. Many systems designed to address self-esteem issues aim to provide a "self-care" environment and are intended for individual use. However, self-esteem is known to be influenced by social connections in the real world and personal living conditions [10, 11]. Therefore, collaboration with close associates could potentially improve individuals' self-esteem. For instance, mobile applications, such as social media, can foster connections with family and friends. However, the mechanisms for enhancing self-esteem via interactions within mobile applications remain largely unexplored.

1.2 Our Research

Our study aims to develop a self-esteem support system that enables users to exchange "episodes" (memories of personal experiences) that reflect personal strengths with friends and family members. In particular, we discuss the essential functionalities and design frameworks required to construct a system that leverages support from others. The framework proposed in this study facilitates the exchange of messages of praise that include the strengths of friends, associates and family and episodes about those strengths. Hirano's research [12] has shown that being recognized by others as having strengths can lead to new self-awareness and a positive perception of oneself. The proposed framework attempts to provide specific and persuasive praise by using the user's episodes as message content, not just pointing out the strengths. However, to materialize the proposed system, it is necessary to clarify how episodes induced by others affect self-esteem and mood, as well as the functions and points to consider when creating the proposed framework as a system. Therefore, in this paper, we introduce the proposed framework for a system that facilitates the exchange of episodes about strengths, and we report the results of a pilot study to verify its effectiveness. The pilot study investigates how receiving strength-related episodes from close friends influences individuals' moods and self-esteem and identifies the required functionalities of our proposed system. This serves as an initial step toward realizing the proposed system.

In the pilot study, the participants performed a task that simulated a subset of interactions within the proposed concept, thereby evaluating the effectiveness of the core concept of our framework. Based on these results, we discuss the design of interactions in a support system that encourages the daily exchange of messages describing strengths with friends and family members. Our contributions to the HCI domain are twofold.

- We propose a support design based on exchanging strength-related episodes to enhance self-esteem. Our findings demonstrate that feedback in the form of messages describing individuals' strengths can be useful in supporting self-esteem and positive emotions.

- We identify that the diversity of feedback content describing strengths and the quantitative imbalance during message exchange impact the improvement of positive emotions. These findings provide specific considerations for designing support systems aimed at improving self-esteem via user-created feedback on strengths.

2 Our Proposal and Noteworthy Issues

2.1 Our Proposal

In this study, we propose a framework to support self-esteem through an exchange of online messages about episodes that highlight one's strengths. Our system, a smartphone application, provides users with the opportunity to share written messages, including episodes that highlight personal strengths, through text-based online communication (Fig. 1).

Previous research in the psychological field indicates that recalling autobiographical memories reduces stress and improves self-esteem. In particular, recalling positive episodes helps individuals affirm valued aspects of themselves, which leads to improved self-esteem [13]. However, it is often difficult for individuals to recall specific episodes that highlight these valued aspects. Therefore, providing feedback from others about such episodes may be an effective element for improving self-esteem within our framework.

In the field of HCI, several studies have discussed how to enhance interventions by enabling users to share their activities online. For example, one system transforms personal gratitude journaling into a communal activity, allowing users to share their gratitude practices [14]. However, these studies primarily focus on visualizing and sharing individual activities. In other words, gratitude journaling is content intended for the person who wrote it. In contrast to these approaches, our work aims to identify a method for enhancing self-esteem by providing user-related positive feedback from others.

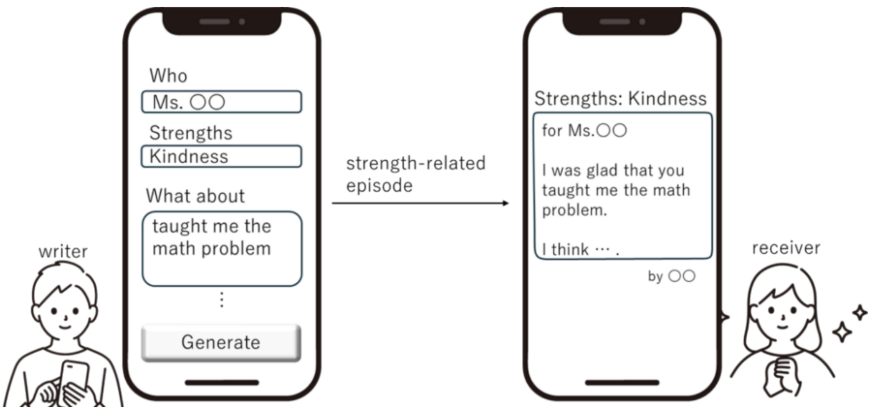


Fig. 1. Overview of our proposed framework.

2.2 Noteworthy Issues

Our approach is expected to help improve self-esteem by emphasizing personal strengths that individuals may not be aware of through feedback from others. However, in order to implement our concept as an effective support system, several issues must be considered.

First, it is not entirely clear whether receiving positive episodes from close friends has a positive effect on mood and self-esteem. The psychological effect of receiving positive feedback about strengths from others has been discussed (e.g., [12]). Nevertheless, while earlier research has focused on the benefits of positive feedback about strengths, it has not addressed how feedback based on personal events might influence psychological outcomes. Therefore, in order to validate the effectiveness of our support concept, it is important to identify how providing feedback based on positive episodes related to the user affects mood and self-esteem.

The second issue is related to clarifying the design of an effective episode-sharing method to support self-esteem through the system. One aspect of this challenge is determining how users input their episodes. For instance, users may find it challenging to record their daily experiences in the system. Although individuals who interact face-to-face frequently have ample opportunities to notice each other's strengths, documenting these observations in detailed written form can be burdensome. Moreover, the way in which the text is shared may influence its effectiveness. In particular, if the feedback or descriptions of personal strengths are overly simplistic, they may lack a sense of authenticity, making it difficult for recipients to recognize their own positive qualities. This may not contribute significantly to enhancing self-esteem. Furthermore, since the proposed system allows users to exchange feedback, factors such as the volume of the feedback received, disparities in feedback between users, and the diversity of the feedback content could also influence users' emotions.

In the next section, as a preliminary step toward designing the proposed system, we present a pilot study aimed at clarifying the effects of positive episodes provided by close friends on mood and self-esteem and what kinds of episodes are effective.

3 Pilot Study

3.1 Purpose

This experiment addresses two issues of clarification that should be considered when implementing the proposed system. The first is related to the impact of receiving messages from others about strengths and episodes related to those strengths on self-esteem and mood. The second is to clarify the necessary functions and considerations involved in creating a system. This pilot study was designed with a focus on the first issue, which underpins our proposed concept. We test the hypothesis that "messages containing positive episodes related to oneself from familiar others are effective in improving self-esteem and mood." As described in detail later (in Sect. 3.2, 3.3), in this experiment, participants emulate the interactions anticipated within the proposed system. Therefore, it is expected that the results will not only confirm the hypothesis but also explore important considerations relevant to the second issue, the design of system interactions.

3.2 Experimental Settings

For this experiment, we recruited pairs of close friends from among Japanese students at our university. The paired participants completed a task of praising each other by exchanging text messages. Table 1 shows the experimental flow. At the beginning of the experiment, the paired participants were separated into different rooms, and each participant was asked to write a message evaluating their partner's strengths. In the message writing task, we instructed them to choose at least two categories from the 24 types of strengths identified by Peterson and Seligman [15], shown in Table 2, and to provide specific explanations as to why they chose them, using the form depicted in Fig. 2 and Table 3.

Table 1. Experimental procedure.

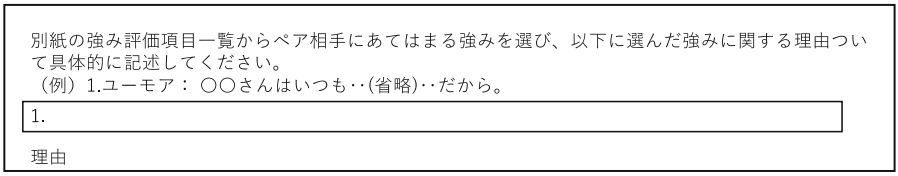
Day		Content	Phase
Day 1	1:	After explaining the experiment and obtaining consent, participants are asked to complete the Rosenberg Self-Esteem Scale	T1
Day 2 (one week after Day 1)	2:	In separate rooms, each participant completes the PANAS to evaluate their mood before the experiment	
	3:	They describe their partner's strengths	Task
	4:	They exchange and review the descriptions	
	5:	They complete the PANAS and the Rosenberg Self-Esteem Scale to evaluate their mood and self-esteem after reviewing the descriptions	T2
	6:	They meet in person with their partner and talk about the descriptions	Task
	7:	They complete the PANAS to evaluate their mood after the experiment	T3

As the participants in this experiment were Japanese, the instructions were presented in Japanese. The English translations of the instructions in the figure are shown in Table 3.

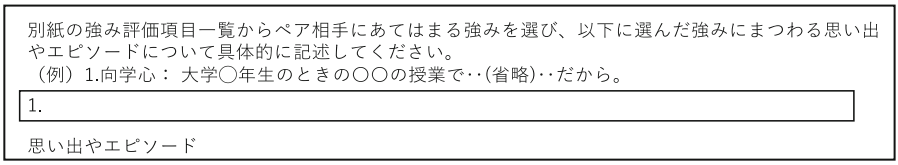
In this task, we set two conditions to verify the effect of episodes when praising strengths. The pairs of participants were assigned to one of two conditions, and they experienced only the same condition together. The first condition (**reason condition**) required participants to explain why they thought their partner had strengths. This condition was intended to lead participants to describe everyday actions that they felt embodied their friend's strengths. The other condition (**episode condition**) required participants to recount a specific episode that highlighted their partner's strengths. This condition was intended to encourage participants to provide detailed episodes in which they felt their partner's strengths were emphasized. After each participant finished writing their message, the two individuals read their partner's message, remaining in their separate room. Next, the pair met back up and were asked to talk about the messages they had

Table 2. The 24 types of strength defined by Peterson and Seligman [15]

1.Creativity	9.Love	17.Prudence
2.Curiosity/love of learning	10.Kindness	18.Moderation
3.Open-mindedness	11.Social intelligence	19.Self-regulation
4.Perspective	12.Citizenship	20.Appreciation of beauty
5.Bravery	13.Fairness	21.Gratitude
6.Persistence	14.Leaderdership	22.Hope
7.Integrity	15.Forgiveness and mercy	23.Humor
8.Vitality	16.Humility and modesty	24.Spirituality



(a) Condition 1: Asking for reasons



(b) Condition 2: Asking for episodes

Fig. 2. Form for filling in the strengths of partners.

Table 3. Experimental conditions, instructions, and examples

Condition 1: Asking for reasons	Condition 2: Asking for episodes
Please select the strength that applies to your partner from the attached word list of strength evaluation items and provide a detailed explanation below as to why you selected it Ex: 1. Humor [Reason] You are always…	Please select the strength that applies to your partner from the attached word list of strength evaluation items and provide a detailed explanation below, including memories and episodes related to the selected strength Ex: 1. Learning [Memory and Episode] In __ class in my __th year of university … So…

received. This session was designed to clarify and correct any misunderstandings about the meaning of the written content.

Table 4. Correlation coefficient r between the number of categories and characters in the messages and the change in PANAS-Positive and -Negative scores between T1 and T2. The symbols indicate the degree of p-values for the test of no correlation. (***: $p < .001$, **: $p < .01$, *: $p < .05$, +: $p < .1$).

	PANAS-Positive (T2-T1)	PANAS-Negative (T2-T1)
Number of categories	**0.338***	–0.012
Number of characters	**0.280 +**	0.228

Table 5. Correlation coefficients and p-values for the relationships between the number of categories and characters in the feedback messages and changes in PANAS-Positive/Negative scores before and after receiving the messages, categorized by self-esteem level. (***: $p < .001$, **: $p < .01$, *: $p < .05$)

Rosenburg Scale (T1)	Number of	PANAS-Positive (T2-T1)	PANAS-Negative (T2-T1)
$\geq$28.58 ($M_{rosenberg}$)	categories	0.258	–0.139
	characters	0.326	0.326
<28.58 ($M_{rosenberg}$)	categories	**0.491***	0.081
	characters	0.194	0.161

In this experiment, participants were asked to complete questionnaires three times: before writing messages, after reading messages, and after the conversation ended.

3.3 Evaluation Metrics

This experiment utilized two questionnaires to measure changes in self-esteem and mood before and after the compliment task and the message conversation task. The first was the Rosenberg Self-Esteem Scale [16], which measures self-esteem; the second was the PANAS (Positive and Negative Affect Schedule) [17] to measure positive and negative moods. Self-esteem was measured twice, before writing the message and after reading it. The PANAS was conducted three times: before writing the message, after reading it, and after the conversation had ended. Table 1 shows the timing of the responses to each questionnaire.

4 Results

Forty university students affiliated with the authors' university participated in this experiment (10 pairs in the reason condition and 10 pairs in the episode condition). The pairs consisted of close friends who interacted regularly with each other. All pairs who participated in the experiment were included in the analysis, except for one pair who provided

incomplete responses in the episode condition. In the following subsections, we compare the results of the effects on mood and self-esteem between the participant group who experienced message exchange under the episode condition and that under the reason condition.

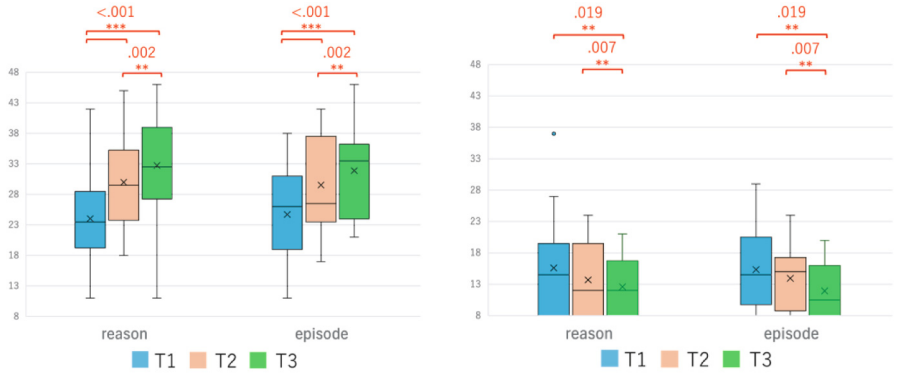
4.1 Does Providing Positive Messages Based on Episodes Help Improve Mood?

Emotional Effects. In the analysis of the PANAS, we calculated positive/negative total scores (PANAS-Positive/PANAS-Negative indices) in T1 to T3, as described in Table 1, and evaluated their differences. Figure 3 presents boxplots of the PANAS-Positive and -Negative scores and the Rosenberg Scale scores for the reason and episode conditions.

Across all experimental conditions, the average PANAS-Positive score increased from T1 to T3, while the average PANAS-Negative score decreased. The results of a two-way repeated-measures ANOVA (Fig. 3), with experimental condition (reason/episodes) and timing (T1/T2/T3) as factors, indicated that the main effect of timing was significant for all measures. However, neither the main effect of the experimental condition nor the interaction between the two factors was significant. Regarding the main effect of timing, multiple comparisons revealed significant differences between all-time points for PANAS-Positive. For PANAS-Negative, significant differences were observed between T1 and T3, as well as between T2 and T3. These findings suggest that being evaluated by others based on one's strengths promotes positive emotions and suppresses negative emotions, regardless of whether the feedback involves reasons or episodes. However, no significant difference was observed between reason-based and episode-based responses.

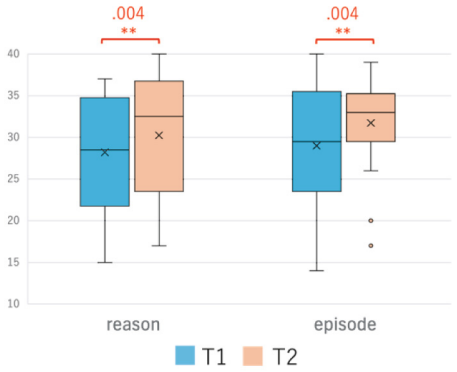
Participants were asked to respond to the Rosenberg Scale only at T1 and T2. Regardless of the experimental condition, the average scores for the Rosenberg Scale at T2 were higher than those at T1 (see Fig. 3). To examine the differences in results between the experimental conditions, we analyzed the results using a two-way repeated-measures ANOVA with experimental condition (reason/episodes) and timing (T1/T2) as factors. Figure 3 shows the results of the ANOVA and post-hoc multiple comparison tests. Similar to the results of PANAS, only the main effect of timing was significant, while neither the main effect of the experimental condition nor the interaction was significant.

These results indicate that receiving feedback from others about one's strengths improves positive mood, reduces negative mood, and enhances self-esteem. Nevertheless, whether the feedback was reason-based or episode-based produced no significant differences in these effects. Consequently, the hypothesis that "presenting specific episodes related to one's strengths evaluated by others increases positive emotions" was not supported. In this experiment, participants were instructed, following the guidelines provided on the response sheet (see Fig. 2 and Table 3), to create feedback that included either reasons or episodes. Additionally, each participant experienced only one of the two conditions (reason or episode). As a result, participants could not be aware of the difference between these two conditions, leading to minimal differences in their descriptions. In some cases, participants wrote episode-like descriptions in the reason condition or reason-like descriptions in the episode condition.



(a) PANAS-Positive

(b) PANAS-Negative



(c) Rosenberg Self-Esteem Scale

Fig. 3. Boxplots and results of the analysis of variance for each indicator (PANAS-Positive, PANAS-Negative, Rosenberg Self-Esteem Scale). The upper figure shows the results of the ANOVA for the timing factor (***: $p < .001$, **: $p < .01$, *: $p < .05$).

Characteristics of Feedback Messages Written by Participants. The boxplots in Fig. 4 illustrate (a) the number of categories and (b) the total length (i.e., the total number of characters) of messages included in each participant response under each condition. In this experiment, no restrictions were imposed on message length, and the only requirement regarding the number of categories was that participants select at least two. As a result, both the number of categories and the length of the feedback messages varied across participants.

4.2 What Kind of Message Affects Emotional States?

From the previous results, our hypothesis that "presenting specific episodes related to one's strengths evaluated by others increases positive emotions" was not supported. However, across both conditions, the results of the PANAS and Rosenberg Scale scores showed significant improvement after the participants received the feedback messages.

These results suggest that our proposed framework, which allows users to receive feedback about their strengths from others, may be effective in improving both mood and self-esteem. Therefore, in order to discuss application design for recording and presenting strength-related messages, we next examine the types of feedback messages that contribute to mood enhancement.

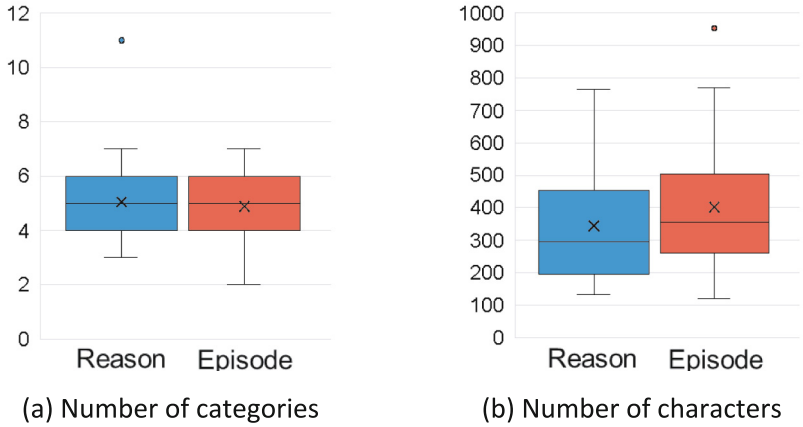


Fig. 4. Numbers of categories and characters contained in participants' feedback messages.

Relationship Between the Content of Feedback Messages and Emotional Changes. As mentioned in the previous section, the content and length of the feedback messages varied across participants. These variations may have influenced the degree of emotional change. To examine factors that may contribute to changes in positive and negative emotions, we analyzed the relationships between the number of categories and characters of feedback messages and the changes in PANAS-Positive and PANAS-Negative scores. Table 4 presents the results of Pearson's correlation coefficients and tests for no correlation. We found weak positive correlations between changes in PANAS-Positive scores and both the number of categories ($r = .338$, $p = .038$) and of characters ($r = .280$, $p = .088$) in the feedback messages. Increases in positive emotions may be related to both the number of included categories and the length of the feedback messages.

To further explore the relationship between the number of categories and emotional change, we divided participants into two groups based on the average number of categories of strengths in their feedback messages (Mcat = 4.97), a "Many" group ($\geq$ Mcat) and a "Few" group ($<$ Mcat). Then, we compared changes in PANAS-Positive scores between the two groups before and after receiving feedback. Figure 5(a) shows the changes in PANAS-Positive scores for both groups. We conducted a two-way repeated measures ANOVA using the number of categories (Many/Few) and the timing of measurements as factors. A significant main effect of timing was detected ($p < .001$). Although the main effect of the number of categories was not significant ($p = .988$), there was a significant interaction between timing and the number of categories ($p = .037$). In

addition, simple main effects tests showed that the increase in PANAS-Positive scores was greater in the "Many" group than in the "Few" group, with differing degrees of significance (Few: $p = .010$; Many: $p < .001$). These results suggest that messages containing a larger number of categories may more effectively enhance positive emotions.

Similarly, we divided participants into "Long" ($\geq$ Mchar) and "Short" ($<$ Mchar) character-count groups based on the average number of characters in their feedback messages (Mchar = 371.55). A two-way repeated measures ANOVA, with the number of characters (Long/Short) and the measurement timing as factors, revealed a significant main effect of timing ($p < .001$). However, neither the main effect of the number of characters nor the interaction effect was significant (see Fig. 5b). Therefore, the influence of message length on recipients' emotional states may be weaker than the influence of the number of categories in the messages.

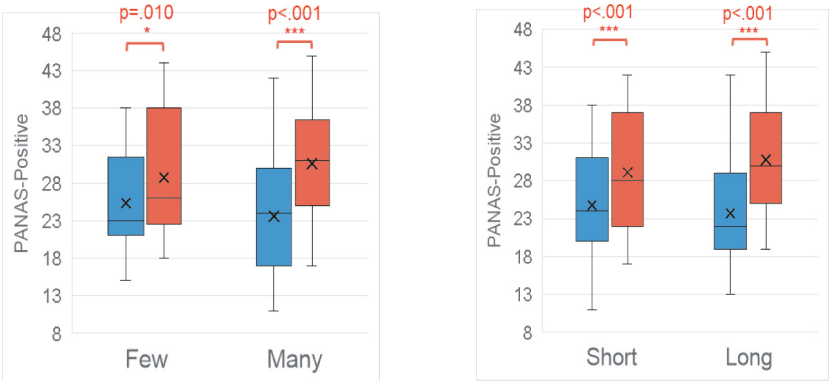


Fig. 5. Changes in results of PANAS-Positive scores before and after receiving feedback messages. (Blue: scores at T1, Red: scores at T2; *: $p < .05$, **: $p < .01$, ***: $p < .001$). Comparison based on the number of (a) categories and (b) characters in the messages.

Moreover, we classified participants into "High" ($\geq$ Mrosenberg) and "Low" ($<$Mrosenberg) self-esteem groups based on the mean score from the Rosenberg Self-Esteem Scale in T1 (Mrosenberg = 28.58). We then examined whether the correlation differed between the "High" and "Low" self-esteem groups. As shown in Table 5, the correlation between changes in PANAS-Positive scores and the number of categories was relatively low in the High self-esteem group ($r = .258$), whereas it was moderate in the Low self-esteem group ($r = .491$). Therefore, feedback messages containing a greater number and variety of categories may be effective at increasing positive emotions, especially among individuals with low self-esteem.

Analysis of Differences in Feedback Between Partners. In this experiment, no upper limits were specified for message length or the number of categories. Therefore, the differences in message length or the number of categories between participants may have

influenced their emotional states. To examine this possibility, we analyzed the correlations between changes in emotional states, as measured by PANAS, and differences in either the number of characters or of categories exchanged between pairs.

The difference was calculated by subtracting the participant's counts from their partner's counts. The results (shown in Table 6) exhibited a weak negative correlation ($r = -.278$) between changes in PANAS-Positive scores and the difference in the number of characters. The significance test approached marginal significance ($p = .091$). In contrast, no correlation was observed between changes in PANAS-Positive scores and the difference in the number of categories. Although the correlation between changes in PANAS-Positive scores and the difference in the number of characters was relatively weak, the results suggest that imbalances in the amount of exchanged messages may affect the enhancement of positive emotional states. To investigate this further, we divided the participants into two groups based on whether the partner's message contained more characters than their own ("Positive" group) or fewer characters ("Negative" group). Table 7 shows the correlations between the differences in the number of characters or categories and the changes in PANAS-Positive/Negative scores for each group. The results of the tests for non-correlation were not significant in any case. However, in the "Negative" group, there was a moderately negative correlation ($r = .372$) between the difference in the number of characters and the PANAS-Positive score. This result indicates that when users write significantly longer messages than their partners, there is a tendency for receiver's positive mood to improve.

Table 6. Correlation coefficients and p-values for the relationships between differences in the number of categories or characters between pairs and changes in PANAS-Positive/Negative scores before and after receiving the messages. (***: $p < .001$, **: $p < .01$, *: $p < .05$).

Difference in the number of	PANAS-Positive (T2-T1)	PANAS-Negative (T2-T1)
Categories	–0.162	0.010
Characters	**–0.278 +**	–0.086

Table 7. Correlation coefficient between differences in the number of categories or characters between pairs and changes in PANAS-Positive/Negative scores before and after receiving the messages, classified by the degree of difference in message length.

Difference in the number of characters	Difference in the number of	PANAS-Positive (T2-T1)	PANAS-Negative (T2-T1)
Partner – Oneself ≥ 0 (Positive group)	Categories	–0.252	–0.089
	Characters	–0.014	0.204
Partner – Oneself < 0 (Negative group)	Categories	–0.049	0.133
	Characters	–0.372	–0.128

5 Discussion

5.1 Outcome of the Experiment

In this section, we summarize the findings from the experiment and discuss the functionalities required in the proposed application to support self-esteem enhancement through interactions with others.

In this experiment, we investigated whether exchanging messages based on episodes related to each other's strengths, which is a core feature of our proposed framework, would improve users' moods and self-esteem. The results showed no significant difference in the psychological state scores when strengths were given with related reasons compared to when they were given as episodes. In any case, these scores increased significantly between before (T1) and after (T2) receiving the feedback message. Therefore, the concept of exchanging messages in pairs, in which individuals praise each other's strengths, may be effective in supporting an improvement in emotional state.

When we explored cases in which PANAS scores increased, we found a slight correlation with the number of categories in the feedback messages. In particular, participants with low self-esteem tended to show a stronger correlation with mood changes than those with high self-esteem. These results suggest that providing feedback from a variety of perspectives may be important for improving mood. In addition, imbalances in the length of feedback messages exchanged between pairs may influence the degree of enhancement of mood.

5.2 Future Directions for the Proposed Framework Based on the Results

Based on these findings, two points can be incorporated into the design of our proposed system, which aims to enhance self-esteem through the exchange of positive feedback. The first is to encourage users to adopt a more multifaceted approach when evaluating strengths. If our proposed approach is implemented as an online platform, the system could log past feedback trends and guide users to write about strengths from different perspectives. The second point is to implement methods that adjust differences in message length between users. Large-scale language models, which have been utilized for text summarization and expansion, may be employed to reduce discrepancies in message length.

5.3 Limitations

In this experiment, we created an environment that mimicked the usage of the proposed system. Therefore, it is possible that different effects will be observed during the exchange of online messages through the system. Facilitating future use of the implemented system will require more participants and verification of the effects of long-term use.

6 Conclusion

In this study, we proposed a framework that allows people to exchange messages containing strengths and related episodes with others. We conducted an experiment to clarify two issues that must be resolved in order to realize the proposed framework, a) the impact of episodes given by others on self-esteem and mood and b) the specific functions required of the framework. In the experiment, two comparison conditions were set to verify the effect of episodes. The experiment involved pairs of participants performing the task of writing messages praising each other's strengths. This task was carried out in a situation in which the participants moved to different rooms and did not face each other, as it was assumed that the proposed system would be used in such an environment. The results of the experiment showed that the task of praising others had a positive effect on self-esteem and mood. However, messages containing episodes did not have a significant effect on these benefits. Regarding the content of the messages, it was shown that receiving praise from multiple perspectives and writing messages with a large number of characters had a positive effect on self-esteem and mood.

In future work, we will verify that similar results are observed in experiments within a system that implements our proposed concept. Moreover, we will consider ways to facilitate the use of the system by both the message writer and the message recipient.

Acknowledgments. This work was supported by JSPS KAKENHI Grant Number 23K16931 and Ritsumeikan Global Innovation Research Organization, Ritsumeikan University.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Vogel, E.A., Roberts, L.R., Eckles, K.: Social comparison, social media, and self-esteem. Psychol. Pop. Media Cult. **3**(4), 206–222 (2014)
2. Sowislo, J.F., Orth, U.: Does low self-esteem predict depression and anxiety? a meta-analysis of longitudinal studies. Psychol. Bull. **139**(1), 213–240 (2013)
3. Abouserie, R.: Sources and levels of stress in relation to locus of control and self-esteem in university students. Educ. Psychol. **14**(3), 323–330 (1994)
4. Inkster, B., Sarda, S., Subramanian, V.: An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. JMIR Mhealth Uhealth **6**(11), e12106 (2018)
5. Prochaska, J.J., et al.: A therapeutic relational agent for reducing problematic substance use (Woebot): development and usability study. J. Med. Internet Res. **23**(3), e24850 (2021)
6. Sharma, A., Rushton, K., Lin, I. W., Nguyen, T., Althoff, T.: Facilitating self-guided mental health interventions through human-language model interaction: a case study of cognitive restructuring. In: Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI'24), pp. 1–29 (2024)
7. Nepal, S., et al.: MindScape study: integrating LLM and behavioral sensing for personalized AI-driven journaling experiences. In: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (UIST'24), vol. 8, no. 4, pp. 1–44 (2024)

8. Tahir, H.M., Richards, D., Kangas, M.: The use of an embodied conversational agent to improve self-esteem: an initial study. In: ACIS 2024 Proceedings, vol. 128, pp. 1–15 (2024)
9. Bhattacharjee, A., et al.: Actually I can count my blessings: user-centered design of an application to promote gratitude among young adults. Proc. ACM Hum.-Comput. Interaction **8**(CSCW2), 1–29 (2024)
10. Muriana, L.M., Baranauskas, M.C.C.: Technological influence on self-esteem: towards a research agenda through a systematic literature review. In Proceedings of the 23rd International Conference on Human-Computer Interaction (HCI'21), LNCS, vol. 12762, pp. 451–471 (2021)
11. Lopez, S., Snyder, C.: Positive Psychological Assessment: A Handbook of Models and Measures. American Psychological Association, Washington, DC (2003)
12. Hirano, M.: Self-praise through "praising others" and "being praised by others": toward effective resilience education for Japanese female college students. Bullet. Tokyo Kasei Univ. 1, Cult. Soc. Sci. **59**(1), 61–70 (2019). (in Japanese)
13. Wildschut, T., Sedikides, C., Arndt, J., Routledge, C.: Nostalgia: content, triggers, functions. J. Pers. Soc. Psychol. **91**(5), 975–993 (2006)
14. Ho, B., Lei, K., He, J.X., Itakura, R., Lum, K., Lee, D.: A pilot study on people's views of gratitude practices and reactions to expressing gratitude in an online community. In: Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing (CSCW'23 Companion), pp. 182–188 (2023)
15. Peterson, C., Seligman, M.E.P.: Character Strengths and Virtues: A Classification and Handbook. Oxford University Press/American Psychological Association, Washington, DC (2004)
16. Mimura, C., Griffiths, P.: A Japanese version of the rosenberg self-esteem scale: translation and equivalence assessment. J. Psychosom. Res. **62**(5), 589–594 (2006)
17. Yamasaki, K., Katsuma, R., Sasaki, A.: Development of a Japanese version of the positive and negative affect schedule for children. Psychol. Rep. **99**(2), 535–546 (2006). (in Japanese)

Telling the Truth: AI-Driven Optimization of Online Communication

Xinyue Xu(✉) and Ruonan Zheng

Tsinghua University, Beijing, China
xuxinyue23@mails.tsinghua.edu.cn

Abstract. Maintaining interpersonal relationships is a major challenge in modern communication. When communicators coordinate information content and the relationship between the two parties, they often face a sense of imbalance caused by dimensional misalignment. This paper proposes and verifies a theoretical model to reveal how LLM can intervene in online communication through strategic AI, thereby regulating users' perception of rapport and message accuracy. We proposed the concept of a "Tell the Truth" plug-in, which aims to help users deal with the multi-dimensional relationships of information in online communication environments through strategic AI intervention. We conducted an online experiment with 97 participants. They completed three rounds of online response tasks based on work situations. We manipulated the intervention strategy of AI in communication (no AI, strategic AI, non-strategic AI). Results showed that AI intervention improved the semantic accuracy of message texts and affected participants' perception of interpersonal rapport. This effect was mediated by cognitive load and moderated by task difficulty. Quantitative evidence suggests that strategic AI intervention in online communication is a powerful means to maintain interpersonal relationships. This study provides a new perspective for understanding the information processing mechanism of AI-MC. It applies the concepts of human-computer interaction (HCI) and human-human interaction (HHI) to study the dynamics of interpersonal communication under AI intervention. It provides theoretical guidance and practical inspiration for developers and designers.

Keywords: AI-Mediated Communication · Interpersonal Dynamics · Information Processing

1 Introduction

Interpersonal communication is crucial for us to maintain social relationships. Effective communication not only enhances the depth and quality of interpersonal relationships but also serves as the main channel for personal emotional expression and disclosure [3,41]. As computer-mediated communication (CMC) has been widely used in interpersonal communication for social purposes, online

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 375–394, 2025.
https://doi.org/10.1007/978-3-031-93539-8_25

platforms have also become the main means of information exchange and interpersonal interaction across time and space [28]. However, the relationship sets in online communication environments often show highly heterogeneous characteristics [21] which means that people often face some challenges when maintaining interpersonal relationships online. As Knapp pointed out, each message between two people carries two layers of information: the content layer and the relationship layer [41]. However, in many social relationships, there may be a misalignment between the content dimension and the relationship dimension of the message [58], which is manifested in that the specific information or facts conveyed have a more complex modal relationship with the social relationship status between the sender and the receiver. Sprecher(2014) proposed that when interacting through the Internet, people will feel uncomfortable with the imbalance caused by non-mutual disclosure within the synchronous relationship form (such as between strangers) [57]. This type of asynchronous communication often occurs and significantly affects interpersonal trust [12,21]. Therefore, communicators often face challenges in coordinating the content of the information and the relationship between the interlocutors (Fig. 1).

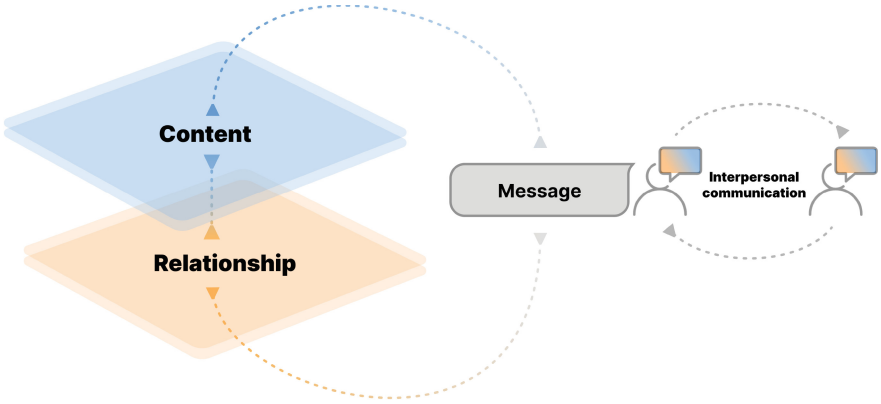


Fig. 1. The two-layer dimension of messages in interpersonal communication.

In recent years, with the development of ChatGPT and other technologies supported by large language models (LLM), artificial intelligence-mediated communication (AI-MC) has made positive efforts to help users communicate effectively by modifying, enhancing, and even generating communication goals through computational agents [22,25]. Unlike computer-mediated communication (CMC), which acts as a passive medium, AI-MC enables AI to play an active role and shape the communicator's goals and actions [11,32]. AI-MC tools are becoming more and more involved in the interpersonal communication process, and there is growing evidence that artificial intelligence has a huge impact on communication frameworks [48]. For example, LLM tools supported by companies such as OpenAI and Anthropic have begun to be widely used around the

world. Grammarly, a cloud-based AI writing assistant, has more than 40 million people and 50,000 organizations using as their published AI writing partners as of 2024 [22]. In the CMC environment, the intervention of AI has become an emerging communication paradigm [32].

Although studies have shown that AIMC tools can improve communication efficiency and positivity, some scholars have questioned that current research focuses mainly on technology and ignores how AI-generated messages affect interpersonal relationships at the social level [33]. People also believe that the mere existence of smart replies will change the way they communicate, partly because of the language bias of smart replies, which generally tend to convey more positive emotions [31]. In addition, although smart replies can help enhance communication skills, they may also have negative effects, such as a crisis of trust [35] and an unfavorable view of others [33,64]. These studies show that how AI intervention affects the transmission of interpersonal relationships in the process of information transmission is still a topic that needs to be explored in depth.

Our current study addresses these gaps by examining how to intervene in online communication processes with LLM to help users process the multidimensional relationships of information and develop a conceptual model to explain the underlying process. To achieve this goal, we designed an online reply process to evaluate the impact of different reply strategies on interpersonal communication. Our research goal is to measure how participants perceive the rapport and accuracy of their messages when AI intervenes in the interaction process to different degrees The results show that AI intervention can effectively affect interpersonal perceptions, and this effect is moderated and mediated by cognitive load. In addition, we find that when AI intervenes in interpersonal communication, the semantic alignment of messages remains consistent and the content expression is clearer. These findings suggest that AI intervention strategies may affect interpersonal relationships and information transfer in online communication environments. This paper provides a new perspective for applying relevant concepts from human-human interaction (HHI) and human-computer interaction (HCI) to understand information processing in AI -mediated online communication.

2 Theoretical Background

2.1 Communication in Shaping Relationships

Communication plays a vital role in interpersonal relationships. It not only affects an individual's perception of others [45,46], and also plays an important role in establishing and maintaining social relationships [49] and achieving cooperative results [66]. Among them, dialogue plays an important role in this process and is a popular medium for maintaining social relationships and coordinating work tasks [5]. Among various forms of communication, interactive dialogue plays an important role as the main way to maintain social connections and coordinate tasks. Past studies have applied conversation analysis to a variety of fields, where researchers have conducted conversation analysis to identify the personality traits and identity information of speakers [45], including using

personality dimensions to identify conversations between leaders [30] and suspected terrorists [61], and dating platforms analyzing text messages to perform personality analysis and increase the chances of successful matching [15].

The effectiveness of communication also varies depending on the type of interpersonal relationship [5]. Research by Duck et al. shows that the quality, purpose, and emotional and social values of communication are profoundly affected by the type of interpersonal relationship [16,19], leading to different communication effects. In addition to semantic content, the text also conveys information about the speaker [45], and this information generally reflects the emotional expression of both parties in the conversation [63], the adjustment of social relations, and the confirmation of social roles. This is consistent with the two-dimensional communication theory proposed by Knapp [41], which shows that the quality of communication depends not only on the transmission of information content but also on the emotional and relational factors in the interaction process.

In the field of computer-mediated communication (CMC), relevant theories believe that communication effectiveness is determined by the interaction between communication needs and media capabilities [14,47]. When the capabilities of the media do not match the communication needs, the content of the communication and the relationship level are often not aligned. Therefore, when improving communication effectiveness, we believe that the adjustment of the media itself is also particularly important [26]. Artificial intelligence-mediated communication (AI-MC), a form of communication in which artificial intelligence modifies, enhances, or generates information during the communication process [25], has also received widespread attention in recent years. Research shows that AI can significantly improve the effectiveness of communication by changing language structure and content generation methods [48]. For example, generative language models have been shown to improve the quality of political conversations by helping users speak to their political opponents in a more respectful and accepting manner [4]. J Chia and others found that AI can influence individuals' views on social issues [13] by adjusting their language style.

Therefore, this study aims to explore how AI-MC systems can improve the quality of interpersonal communication by optimizing the coordination between the content and relationship dimensions of information and to provide new perspectives and solutions for the theory and practical application of information transfer.

2.2 Cognitive Load in Communication and Technical Support

In recent years, the issue of information overload has received widespread attention from the academic community [17]. Computer-mediated communication systems (CMCS) have greatly enhanced the convenience of information exchange and broken through the limitations of time and space [28]. However, this convenience also brings some negative effects. CMCS enables individuals to experience higher communication stimulation by significantly increasing "social density" and "connectivity" [8]. Klapp points out that excessive information flow can lead to distraction, increased stress, and more errors, resulting in poorer information [40]. Kerr also pointed out that the amount and speed of information can be overwhelming, especially when the messages are not linear and often contain multiple threads [36,62].

Research shows that by optimizing certain structural designs of the interaction space in CMCS, information processing can be effectively promoted, thereby reducing information overload and the pressure it brings [29]. Therefore, understanding how technology supports cognitive processing in the communication process is of great significance for improving information exchange in complex environments. Artificial intelligence-mediated communication (AI-MC) can effectively adjust the way information is presented and improve communication efficiency [25]. Based on the above insights, this study proposes the following hypothesis:

H1: In online communication, AI-mediated interactions are perceived to involve less cognitive load.

The level of cognitive load is generally considered to be related to the complexity of the task, especially the complexity of language tasks usually leads to a decline in individual performance [39]. The communication process intervened by AI can effectively affect the way of communication and interaction. [51,59], and different types of AI intervention strategies can also effectively change communication patterns. Based on the above insights, this study proposes the feasibility of reducing communication difficulty through AI intervention to reduce cognitive load, and proposes the following hypotheses:

H2: Communication strategies will affect the mobilization of individual cognitive resources.

Specifically, this change was mediated by task difficulty, and difficulty perception was linearly related to cognitive load.

However, as the availability of AI technology increases, its potential negative effects are also gradually emerging. People believe that texts marked as AI-generated are untrustworthy and that AIMC cannot accurately convey their intended information [44]. However, some empirical studies have shown that AI-generated texts can reduce content ambiguity through semantic consistency so that users do not need to be distracted by repeatedly proofreading content when processing information [9]. Currently, many companies have tried to integrate AI into text creativity to support various writing tasks [22]. Based on previous research on AI semantic similarity, we hypothesize that:

H3: In a communication environment, AI can ensure consistency in the semantic content delivered to individuals.

2.3 Cognitive Load and Relationship

Communication is the basis for building trust and rapport between two parties. Appropriate manners and positive interactions are key elements of positive rapport, promoting the building of positive interpersonal relationships [10]. In face-to-face communication, these factors are easily communicated and interpreted through direct verbal and nonverbal signals. However, online communication is different from face-to-face communication, and most communication needs to be conducted through text [37]. Participants cannot see or hear each other, and communication is often asynchronous, with delays between responses, in a pattern similar to the exchange of written communications [54]. These characteristics may significantly affect the formation of conversational feedback [How Awareness of Each Other's Mental Load Affects Dialogue] and, therefore, the establishment of rapport between the two parties in the conversation. Especially in situations where visual and non-verbal cues are lacking, individuals need to invest more cognitive resources to fill this information gap to effectively establish rapport [18]. Therefore, extra attention needs to be paid to the management of cognitive load in online communication environments to reduce the risk of conflict. Based on this background, this study proposes the following hypotheses:

H4: During online communication, the consumption of cognitive resources has a significant impact on the maintenance of relationships between individuals.

there is a linear correlation between cognitive load and perceived affinity.

Finally, combining Hypothesis 1, Hypothesis 2, and Hypothesis 4, we further propose the following hypothesis:

H5: Response strategies can affect the perception of interpersonal relationships.

Specifically, reply strategy was strongly associated with the construction of rapport with others and was moderated by cognitive load.

Overall, we proposed a moderated mediation model (Fig. 2) that integrates Hypotheses 1, 2, and 4 to understand the process of change in interpersonal perceived affinity in AI-MC. The model takes into account the following ideas: The process of AI intervention in interpersonal communication may change the individual's cognitive load level (H1), and different response strategies will affect cognitive load, which may be regulated by the difficulty of the response task. (H2). Furthermore, it acknowledges that cognitive load may have a significant impact on interpersonal rapport, and the two are linearly related (H4). Finally, the model indicates that communication strategies affect interpersonal rapport through the influence of cognitive load (H5).

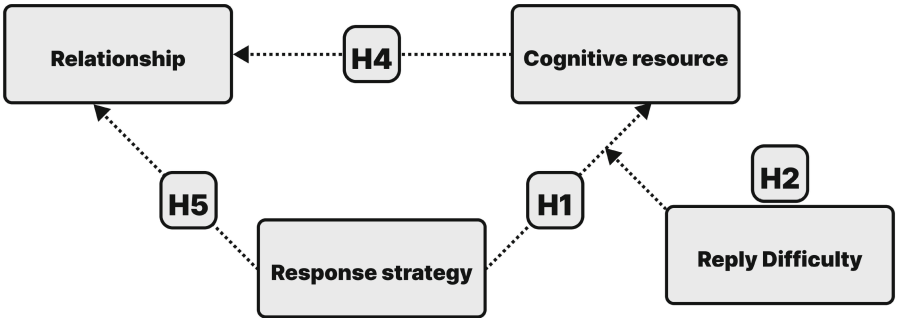


Fig. 2. The Proposed Moderated Mediation Model of AI-Mediated Communication

3 Methods and Materials

3.1 Online Work Communication

In the study of online communication relationships, many experimental scenarios are based on work environments. These online work platforms have become key places for studying trust, interpersonal relationships, and information exchange [1]. Interactions in these environments can effectively reveal the role of individuals in the group, information flow, and collaboration methods, and provide the possibility of parameter adjustment of emotional stimulation. Through these studies, scholars can deeply analyze how individuals influence interpersonal perception and information flow through text, voice, video, etc. without face-to-face interaction. In addition, some scholars believe that workplace friendships and organizational citizenship behavior (OCB) are key concepts for understanding employee dynamics [43] and are effective tools to promote understanding of the complexity and diversity of interpersonal relationships. Therefore, we situate our work in a daily online work communication scenario.

Participants are asked to imagine that they are doing a part-time job and receive a text message from the leader hoping to work overtime on weekends, but this request conflicts with the participant's weekend schedule. Participants need to reply according to our requirements. In our study, we set the reply into three rounds, namely, user self-edited text reply, strategic AI participation reply, and pure AI reply. The purpose is to let users perceive the interaction with different reply strategies in the same situation, to perceive the reply load and rapport under different behavior patterns.

3.2 Strategic AI

the strategic intervention of AI in the interpersonal online communication environment, we designed an AI reply strategy based on the research of Knapp, dividing each message into two dimensions: the content layer and relationship layer, to study the impact of strategic AI on reply difficulty, maintenance of interpersonal relationships and message content. In the experiment, we designed three AI intervention strategies, adjusting their preferences for identity perception and content optimization respectively, so that participants can perceive the communication performance of different strategies in the AIMC environment, which are:

- **Partial Strategy AI1 (No identity awareness + Accurate content):** AI only generates responses based on a general language model, without social adjustment strategies, but expresses accurate message content.
- **Partial strategy AI2 (identity awareness + no content accuracy):** Adjust the language style according to social relationships (such as superiors, colleagues, friends) to ensure the appropriateness of the relationship, but do not require accuracy at the message content level.
- **Full-strategy AI (identity awareness + content accuracy):** AI considers both social identity and content optimization to provide responses that are in line with social context and clearly expressed.

Through this non-strategic AI intervention, we hope to explore how AI changes users' communication decisions and social interaction effects under different combinations of identity awareness and content optimization.

3.3 Research Design

This study adopted a single-factor within-subjects design, in which the only independent variable was the response method (participant self-editing, with strategic AI intervention, and without strategic AI intervention). In the experiment, each participant completed the three experimental conditions in a counterbalanced order, and all participants had to experience all experimental conditions.

We published online questionnaires through social media platforms and collected a total of 155 participant data. To ensure data quality, we screened the questionnaires and eliminated those that were incomplete, had abnormal reaction times, and finally obtained 97 valid data. The age distribution of the participants ranged from 19 to 63 years old ($M = 26.30$, $SD = 7.33$) All data are strictly confidential and will only be used for academic research.

3.4 Procedure

The experiment was conducted by combining online experiments with questionnaire measurement. The participants' task was to reply to messages in three different ways (self-editing, strategic AI collaboration, and non-strategic AI collaboration) according to the task requirements. After each reply, participants evaluated the data of each dimension of the reply process. The design of this experiment refers to Hohenstein et al. (2023) research on how AI affects interpersonal interaction and perception [33], and an online survey platform was used for data collection to ensure the reproducibility of the experiment and the reliability of the data. The detailed process of the experiment is shown in Fig. 3.

After entering the online experimental environment, all participants were required to read the experimental instructions and sign the informed consent form online. The experiment simulated an online communication situation. Participants were asked to imagine being in a work scenario. They were asked to respond after receiving a work request from their leader. The request involved an arrangement for overtime work on the weekend, but it conflicted with the participant's plan. The message received was as follows:

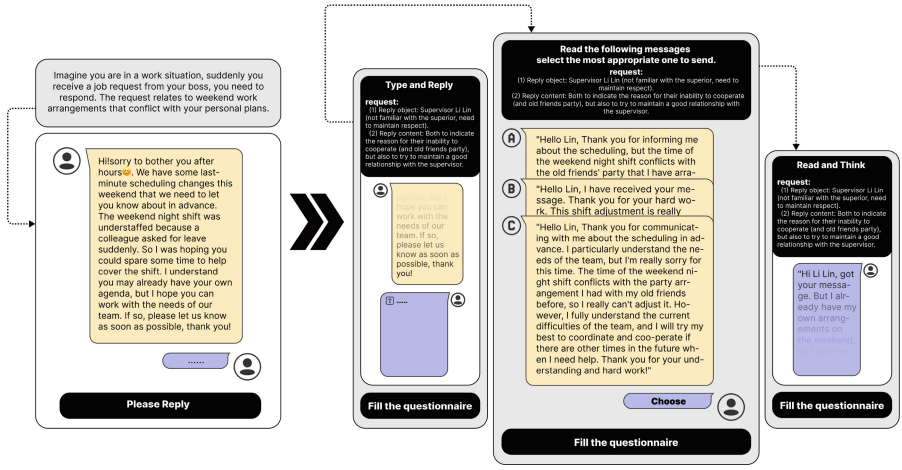


Fig. 3. Experimental Procedure.

"Hi, sorry to bother you after work. We have some temporary adjustments to the schedule this weekend, so we need to tell you in advance. Because a colleague suddenly asked for leave, the weekend night shift is short of staff. So I hope you can spare some time to help take over this shift. I understand that you may have your arrangements, but I also hope that you can cooperate with the needs of our team. If it is not possible, please let us know as soon as possible, thank you!"

Next, participants need to complete three reply tasks in sequence according to the prompts. The task requires that the reply first needs to maintain a harmonious relationship with the leader, and then needs to convey the reason for the refusal (getting together with old friends). The three reply strategies are as follows:

- **Self-edited reply:**
  Participants need to edit the reply text by themselves, without AI intervention.
- **Partial strategy AI intervention:**
  The system provides 3 alternative responses generated by AI (1. No identity awareness but accurate content; 2. Identity awareness but inaccurate content; 3. Identity awareness and accurate content), and participants need to choose 1 to send.
- **No-strategy AI task:**
  The system automatically generates AI responses that participants cannot edit or select.

After each response task, users were asked to fill out a questionnaire, which mainly included the evaluation of task load, cognitive load, self-perceived harmony, and semantic similarity.

3.5 Date Analysis

Since this experiment involves repeated measurement data and hierarchical structure data, the linear mixed effects model (LMM) was used for statistical analysis, and the

specific implementation was implemented using the lmertest package in the R language [42]. Among them, because the variables involved selection situations and the data distribution was unbalanced under the condition of strategic AI intervention, we used the weighted regression model (Weighted Regression Model) for adjustment to reduce the impact of imbalance on the results [23].

Descriptive statistical analysis was first performed in the data analysis process, and zero-order bivariate correlations were calculated to assess the basic relationships between the main study variables. Subsequently, the hypothesized moderated mediation model (Process Model 14, [27]) was tested in a single model framework using Bootstrap (n = 5000, CI = 95%) to assess the significance of the indirect effect of the moderator at different levels [50]. At the same time, this study uses the pre-trained cross-language model provided by the Sentence-Transformers library [53] to analyze the differences in message text accuracy of the cross-language model under different communication strategies.

This study first conducted a moderated mediation analysis on interpersonal perceived affinity (Rapport). Among them, different communication strategies were used as predictors, cognitive load under different communication strategies was used as a mediator, task difficulty was used as a moderator, and interpersonal perceived affinity (Rapport) was used as the dependent variable for analysis [6]. Then, the accuracy of message texts under different communication strategies was analyzed using a cross-language model, and an independent sample T-test was performed. The above analysis was completed in the R language environment, lavaan was used to estimate the structural equation model, and lme4 was used to fit the mixed linear model [7].

4 Results

Table 1 lists the descriptive statistics of all variables, including means, standard deviations, and zero-order correlations. In addition, Fig. 4 A shows the participants' perceived harmony scores under different response strategies. The results of one-way analysis of variance (ANOVA) showed that the data met the homogeneity of variance (Levene test, $p > 0.05$), and the perceived harmony under strategic AI intervention was significantly higher than that under no AI intervention condition (none-AI: $F = 2.24$, $p < .00\ 5$; Strategic AI: $F = 1.76$, $p < .005$). However, no statistically significant difference was observed in the Non-strategic AI (NSA) intervention condition (AI: $T = 0.746$, $P = .664$). Figure 4 B shows the cognitive load and task load scores under different response strategies. To evaluate the internal consistency of the measurement tools, we calculated Cronbach's α coefficient, and the results showed that task load ($\alpha = 0.81$), cognitive load ($\alpha = 0.94$), and perceived rapport ($\alpha = 0.91$) all had high internal consistency and good reliability.

The results of the moderated mediation analysis are shown in Table 2 and Fig. 5. The analysis used the PROCESS method of Hayes [27], and confidence intervals were estimated by 5000 bootstrap sampling. The results showed that the response strategy significantly affected cognitive load (path a: $\beta = -0.43$, $z = -4.42$, $p < 0.001$, supporting H1), the effect of cognitive load on perceived interpersonal harmony (path b: $\beta = 0.26$, $z = 4.52$, $p < 0.001$, supporting H4), and the direct effect of response strategy on perceived harmony was also significant (path c: $\beta = -0.37$, $z = -4.39$, $p < 0.001$), but was partially mediated by cognitive load (supporting H5). The mediation effect test showed that the direct effect was $\beta = -0.37$, 95% CI $[-0.19, -0.05]$, $p < .01$, and the indirect effect was $\beta = -0.11$, 95% CI $[-0.54, -0.21]$, $p < .001$, indicating

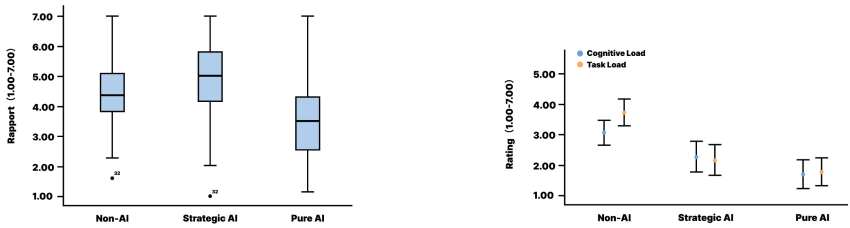
Table 1. Means, standard deviations, and zero-order correlations of the investigated variables.

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Age	26.33	7.32	–													
2. Gender	1.77	.42	−0.08	–												
3. NATL	4.23	1.38	−0.10	.00	–											
4. NACL	3.82	1.29	.00	−0.01	.68**	–										
5. th NAR	4.46	1.03	−0.02	.04	−0.03	.08	–									
6. NASS	.70	.11	−0.01	−0.15	.03	−0.01	−0.14	–								
7. SATL	3.27	1.57	.02	.26**	.23*	.44**	.18	−0.28**	–							
8. SACL	3.34	1.58	.01	.24*	.22*	.45**	.24*	−0.14	.92**	–						
9. SAR	4.95	1.18	.04	.08	.13	.15	.61**	−0.03	.20	.25*	–					
10. SASS	.74	.07	−0.25*	−0.13	−0.01	−0.03	.06	.17	−0.09	−0.01	.09	–				
11. PATL	3.01	1.42	−0.07	.16	.30**	.49**	.11	−0.31**	.67**	.63**	.17	.04	–			
12. PACL	2.95	1.47	−0.07	.15	.32**	.47**	.12	−0.23*	.72**	.71**	.18	.04	.89**	–		
13. PAR	3.49	1.29	−0.03	.05	.05	.24*	.22*	−0.20	.39**	.38**	.255*	.03	.52**	.56**	–	
14. PASS	.79	0	.b	.b	.b	.b	.b	.b	.b	.b	.b	.b	.b	.b	.b	–

Note: N = 97. Gender: 0 = female, 1 = male. NATL = Non-AI Task Load; NACL = Non-AI Cognitive Load; NAR = Non-AI Rapport; NASS = Non-AI Semantic Similarity; SATL = Strategic AI Task Load; SACL = Strategic AI Cognitive Load; SAR = Strategic AI Rapport; SASS =Strategic AI Semantic Similarity; PATL = Pure AI Task Load; PACL = Pure AI Cognitive Load; PAR = Pure AI Rapport; PASS = Pure AI Semantic Similarity; *p < .05. **p < .01. b Unable to calculate because at least one variable is a constant.

that the model conformed to the partial mediation model. In addition, task load has a significant moderating effect in the path between cognitive load and perceived rapport (supporting H2), and cognitive load is significantly positively correlated with task load. The moderated mediation index (Index of Moderated Mediation) of the overall moderated mediation model was 0.07 4, $p < 0.001$ (Hayes, 2018), and the Bootstrap confidence interval did not cross 0, further verifying the establishment of the moderated mediation effect.

Finally, we examined the semantic similarity between the message text and the task requirements under different communication strategy conditions. An independent sample T-test was performed on the message texts under the conditions of no AI intervention (None-AI) and strategic AI intervention (Strategic AI). The analysis results showed that there was a significant difference in the semantic similarity between the two groups of texts ($t = -2.77$, $p<.01$), and the mean difference was negative, indicating that the semantic similarity without AI intervention was significantly lower than that with strategic AI intervention (95% CI = $[-0.06, -0.01]$). In addition, the semantic similarity in the pure AI generation (Pure AI) condition was constant (mean m = 0.79), which was significantly higher than that in the no AI and strategic AI groups ($p < 0.01$), further supporting H3.



(a) (b)

Fig. 4. A shows the data distribution and median of the participants' perceived rapport ratings (n = 97). B shows the mean and 95% confidence interval of the ratings given by the participants on the level of load (n = 97).

Table 2. The Moderated Mediation Effect of Response Strategy on Rapport (via Cognitive Load, with Response Strategy as Moderator)

Path	Estimate (β)	Std. Err	z/t-value	p-value
Model 1 (Mediation) – CL as Mediator				
RS → CL (a)	−0.43	0.10	−4.42	<0.001
CL → R (b)	0.26	0.06	4.52	<0.001
RS → R (c')	−0.37	0.09	−4.39	<0.001
Indirect Effect (a×b)	−0.11	0.03	−3.22	0.001
Direct Effect (c')	−0.37	0.09	−4.39	<0.001
Model 2 (Moderation) – R as Outcome				
RS → R	−1.37	0.20	13.09	<0.001
CL → R	−0.36	0.13	−2.91	0.004
RS × CL	0.29	0.06	5.24	<0.001

Note: N = 97. All data are standardized. RS: Response Strategy; R: Rapport; CL: Cognitive Load.

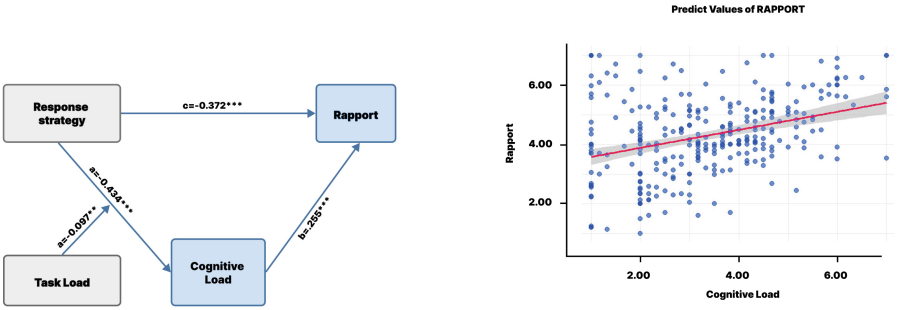


Fig. 5. (a) Shows the results of the moderated mediation analysis on the relationship between perceived rapport, response strategy, cognitive load, and task load. The values represent the standardized regression coefficients of the response linear mixed model ($*$ $p < .05$). (b) This figure shows the marginal effect of cognitive load on the perceived rapport dimension, with the points being the standardized sample data ($N = 97$), the regression line being the estimated effect, and the shaded area representing the 95% confidence interval.

5 Discussion

In summary, this study explored how different degrees of AI intervention affect interpersonal perceived affinity in an AI-MC environment related to interpersonal communication. Combined with the analysis of message semantic similarity, we further examined how the large language model (LLM) intervenes in the online communication process to optimize information processing and thus affect the multidimensional relationships of information processing by individuals. The results of the study show that in the context of strategic AI intervention, individuals' perception of interpersonal affinity is significantly improved. This positive relationship is achieved by AI reducing the difficulty of the response task and thus adjusting the individual's cognitive load.

5.1 AIMC and Perceived Rapport

As predicted in Hypothesis 1, our study found that when it comes to online communication scenarios, AI-mediated interactions significantly reduce participants' cognitive resource consumption compared to self-edited responses, indicating that response strategies affect individuals' cognitive load. During the experiment, the response task remained consistent, and no intervention measures were set to directly affect the participants' cognitive perception. We speculate that the variation in cognitive load mainly stems from the participants' response mechanisms to different response strategies. This result is consistent with research in the field of education, such as AI-assisted language learning (AIAL) strategies (personalized feedback, interactive practice, intelligent tutoring), which have been shown to effectively reduce students' cognitive load and improve learning outcomes [20].

The results of mediation analysis showed that there was a linear relationship between individuals' cognitive load and perceived affinity and that the consumption of cognitive resources during online communication had a significant impact on the

maintenance of interpersonal relationships. This is highly relevant to previous theories on interpersonal relationship cognition. This is highly consistent with the theories of previous studies on cognition and social interaction. As Chen's research points out, a high cognitive load will change trust behavior in interpersonal interactions, and affect language selection and pause frequency, and a high load will also inhibit internal behavioral coordination, thereby impairing empathy and collaboration capabilities [56]. A high workload can inhibit the coordination of an individual's internal behavior, which may affect empathy and collaboration in social interactions [34].

Moderation analysis further revealed that the relationship between response strategy and cognitive load was moderated by the difficulty of the response task, and the two were positively correlated. Studies have shown that the AI-mediated response process can significantly reduce individuals' subjective perception of task difficulty. This result echoes findings in the field of human-machine collaborative decision-making, namely that different task complexities have a significant impact on the effectiveness of human-machine collaboration, and emphasizes that AI explanation models should be optimized according to task complexity to improve human performance in complex decision-making [55].

The study also found that response strategies can significantly influence perceptions of interpersonal relationships. Reply strategy was strongly associated with building rapport with others and was moderated by cognitive load. This finding is consistent with previous research on role perception of AI. Previous studies have shown that AI will be given different emotions [38] in different role settings. For example, in the workplace, employees have privacy concerns about AI (such as sentiment analysis and emotion detection tools), which leads to higher emotional defensiveness and affects interpersonal trust in the team [24].

AI-generated texts have higher accuracy in semantic expression than human autonomous responses, especially when using strategic AI models, where semantic similarity is significantly improved. This is consistent with previous research, such as the application of deep learning and natural language processing (NLP) technology in semantic search, which can correct spelling and grammatical errors and improve the relevance of search results [52]. In addition, in the semantic communication task, the AI-driven communication system (SC-AIT) achieved 40% higher accuracy in image classification tasks than traditional methods, further demonstrating the potential of AI in improving the quality of semantic content [65].

Although this study is based on a preset work scenario, we believe that its results have broad real-world applicability. The work environment encompasses important aspects of real-world communication, involving multiple dimensions such as organizational structure, virtual collaboration, task uncertainty, and leadership, which together affect employee satisfaction and social relationships [60]. Therefore, we believe that in this experiment, the scenario created based on the work environment effectively simulates the key features of AI intervention in interpersonal communication in complex communication activities in the real world.

Although the interpersonal perceived rapport in the experiment was triggered by different communication strategies, this dimension represents a general interpersonal relationship pattern rather than a specific perception of a particular communication behavior. The study shows that the rapport derived from communication behavior has a broader interpersonal relationship meaning, not just limited to specific interaction situations [2].

In summary, in the online communication environment, the communication process with AI intervention can bring better interpersonal perception harmony and semantic

content accuracy. Among them, compared with the cases without AI and strategic AI, strategic AI achieved better results in the collaborative performance of the content layer and the relationship layer of the message.

5.2 Implications

This study has important implications for theory and practice. First, this is the first study based on quantitative empirical methods to demonstrate that artificial intelligence (AI) intervention in the communication process can significantly affect interpersonal rapport by regulating cognitive load. The research results not only expand the applicability of AI in daily interpersonal interactions but also provide new empirical support for future research on artificial intelligence and social interaction. In addition, this study combines the theories of HHI (Human-Human Interaction) and HCI (Human-Computer Interaction) to explore how AI-MC affects the formation and development of interpersonal relationships, thereby enriching and expanding the theoretical framework of interpersonal relationship dynamics in the era of artificial intelligence.

The results of the study show that artificial intelligence intervention in online communication processes can effectively shape interpersonal harmony, but at the same time, it may bring unforeseen social consequences by reshaping interpersonal interaction patterns. Therefore, stakeholders (such as social platforms, technology developers, policymakers, etc.) should carefully consider the ethical and social impacts of incorporating AI into social behavior, especially in environments involving large amounts of information load. This study highlights the need for further research in this area and provides stronger academic support for technology governance and artificial intelligence optimization.

5.3 Limitations and Future Work

Some limitations of this study should be acknowledged. First, this study only focused on the harmony dimension of interpersonal relationships, using perceived harmony as the main indicator for measuring interpersonal quality. However, interpersonal relationships are a multi-level and complex concept involving different interaction logics, power structures, cultural values, and cognitive frameworks. Therefore, future research should adopt a more comprehensive perspective to explore the impact of AI- AI-mediated communication on different interpersonal levels, and further analyze the role of AI strategies in different emotional contexts on communication quality. Second, this study was conducted in a specific and controlled environment, which may limit the external validity of the research results. The interpersonal communication environment in reality is dynamic and complex, and the communication behavior of individuals will be affected by specific situations. The experimental environment of this study may not be able to fully capture the behavioral patterns in different communication scenarios, such as conflict communication, cooperative communication, supportive communication, ritual communication, etc. Therefore, future research should conduct experiments in more natural and diverse communication situations to verify whether the conclusions of this study apply to a wider range of communication environments. Third, this study mainly relies on self-report measurement, which may have social desirability bias, that is, the subjects' answers may be affected by self-presentation and fail to fully reflect their real communication experience. Future research should adopt more objective measurement tools. Another limitation is related to the simplification of the AI-MC communication

process. Participants did not interact directly with human subjects but communicated through preset dialogue scenarios. This non-personalized, static dialogue design may not fully simulate actual social interactions, thus affecting the ecological validity of the research results. Therefore, future research should adopt more interactive and adaptive AI-MC systems to make communication closer to real-life situations to improve the external validity of the research.

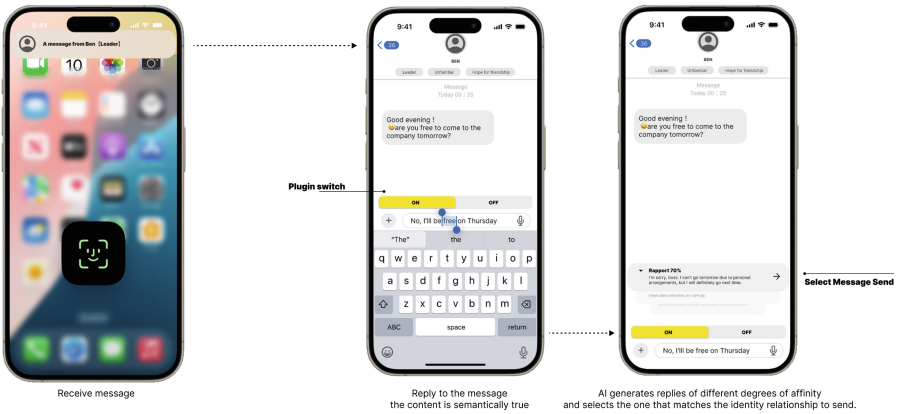


Fig. 6. "Tell the Truth" concept plugin

Future research should explore the impact of independent variables on various forms of interpersonal relationships in a wider range of scenarios. We are developing a communication plug-in called "Tell the Truth" (Fig. 6). We use LLM (GPT-4) to simulate an end-to-end communication assistant to help users process messages and achieve better coordination at the information content layer and the relationship layer. We also use workflow experiments to verify the feasibility of this paradigm.

Finally, the sample size of this study was small and relatively homogeneous, mainly from a college student population. This limitation may affect the generalizability of the research results. Future studies should use larger. Therefore, future studies should expand the sample size and introduce more diverse participants to further verify the robustness and generalizability of the conclusions of this study.

6 Conclusion

the rapid development of AI technology, such as ChatGPT and other language models based on generative pre-trained transformers (GPT), the influence of AI-MC in various social and information interaction scenarios is gradually expanding. However, current research on how to use artificial intelligence-mediated communication (AIMC) technologies to address information processing and social relationship challenges in online communication is still relatively limited. This study aims to explore how different AI intervention strategies affect interpersonal perceptions in online communication and proposes a moderated mediation model to analyze the role of task difficulty and cognitive load in this process. The results of the study show that when AI intervenes

in the communication process, its strategic application can significantly improve individuals' perception of interpersonal rapport. Furthermore, this influence mechanism is mediated by cognitive load and moderated by the difficulty of the response task. The contribution of this study is that it deepens the research on AI-MC in the field of interpersonal communication, provides theoretical support and empirical basis for understanding the role of AI in coordinating information content and interpersonal relationships, and illustrates that different AI intervention strategies can effectively influence interpersonal relationships. These findings provide important references for the future optimization and social application of AI technology and emphasize the need to pay attention to the potential impact on interpersonal relationships when designing AI-assisted communication tools.

Acknowledgments. The authors thank all volunteers who participated in the research related to this publication and all reviewers. This work was supported by the Intelligent Product Innovation Institute, Academy of Fine Arts, Tsinghua University.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Agarwal, S.: The profound impact of virtual communication technologies on team collaboration in the contemporary work environment. Int. J. Embed. Syst. Emerg. Technol. **10**(1), 45–60 (2024)
2. Al-Natour, S., Benbasat, I., Cenfetelli, R.T.: Creating rapport and intimate interactions with online virtual advisors. In: SIGHCI 2007 Proceedings, p. 18 (2007)
3. Altman, I., Taylor, D.A.: Social Penetration: The Development of Interpersonal Relationships. Rinehart & Winston, Holt (1973)
4. Argyle, L.P., et al.: Leveraging AI for democratic discourse: chat interventions can improve online political conversations at scale. Proc. Natl. Acad. Sci. **120**(41), e2311627120 (2023)
5. Avrahami, D., Hudson, S.E.: Communication characteristics of instant messaging: effects and predictions of interpersonal relationships. In: Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work, pp. 505–514 (2006)
6. Baron, R.M., Kenny, D.A.: The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J. Pers. Soc. Psychol. **51**(6), 1173 (1986)
7. Bates, D., et al.: Package 'lme4'. Convergence **12**(1), 2 (2015)
8. Baum, A., Calesnick, L.E., Davis, G.E., Gatchel, R.J.: Individual differences in coping with crowding: stimulus screening and social overload. J. Pers. Soc. Psychol. **43**(4), 821 (1982)
9. Bo, J.Y., Wan, S., Anderson, A.: To rely or not to rely? evaluating interventions for appropriate reliance on large language models. arXiv preprint arXiv:2412.15584 (2024)
10. Braeutigam, A.M.: What i hear you writing is... issues in odr: building trust and rapport in the text-based environment. U. Tol. L. Rev. **38**, 101 (2006)
11. Carr, N.: The shallows: What the Internet is doing to Our Brains. WW Norton & Company (2020)

12. Chen, Q., Zhang, Q., Zhao, S., Li, C.: Trusting strangers: the benefits of reciprocal self-disclosure during online computer-mediated communication and mediating role of interpersonal liking. Int. J. Psychol. **59**(1), 143–154 (2024)
13. Chia, J.: Communicating, connecting and developing social capital for sustainable organisations and their communities. Australasi. J. Reg. Stud. **17**(3), 330–351 (2011)
14. Dennis, A.R., Fuller, R.M., Valacich, J.S.: Media, tasks, and communication processes: a theory of media synchronicity. MIS Q. 575–600 (2008)
15. Donnellan, M.B., Conger, R.D., Bryant, C.M.: The big five and enduring marriages. J. Res. Pers. **38**(5), 481–504 (2004)
16. Duck, S., Rutt, D.J., Hoy, M., Strejc, H.H.: Some evident truths about conversations in everyday relationships all communications are not created equal. Hum. Commun. Res. **18**(2), 228–267 (1991)
17. Edmunds, A., Morris, A.: The problem of information overload in business organisations: a review of the literature. Int. J. Inf. Manage. **20**(1), 17–28 (2000)
18. English, W., Robinson, J., Gott, M.: Health professionals' experiences of rapport during telehealth encounters in community palliative care: an interpretive description study. Palliat. Med. **37**(7), 975–983 (2023)
19. Feldstein, S.: Impression Formation in Dyads: The Temporal Dimension. Human Sciences Press Inc., New York NY, USA (1982)
20. Feng, L.: Investigating the effects of artificial intelligence-assisted language learning strategies on cognitive load and learning outcomes: a comparative study. J. Educ. Comput. Res. **62**(8), 1961–1994 (2025)
21. Finkel, E.J., Eastwick, P.W., Karney, B.R., Reis, H.T., Sprecher, S.: Online dating: a critical analysis from the perspective of psychological science. Psychol. Sci. Public Interest **13**(1), 3–66 (2012)
22. Fu, Y., Foell, S., Xu, X., Hiniker, A.: From text to self: users' perception of AIMC tools on interpersonal communication and self. In: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, pp. 1–17 (2024)
23. Gelman, A.: Struggles with survey weighting and regression modeling. Stat. Sci. 153–164 (2007)
24. Hammad, T.: Exploring the intersection of AI and emotional intelligence: navigating the promise and peril
25. Hancock, J.T., Naaman, M., Levy, K.: Ai-mediated communication: definition, research agenda, and ethical considerations. J. Comput.-Mediat. Commun. **25**(1), 89–100 (2020)
26. Harrison, A., Windeler, J.B.: Framing communication: how agenda alignment and media capabilities shape partially cooperative communication. MIS Q. **44**(2) (2020)
27. Hayes, A.F.: Partial, conditional, and moderated moderated mediation: quantification, inference, and interpretation. Commun. Monogr. **85**(1), 4–40 (2018)
28. Herring, S.C.: Computer-mediated. Ann. Rev. Inf. Sci. Technol. **109** (2002)
29. Hiltz, S.R., Turoff, M.: Structuring computer-mediated communication systems to avoid information overload. Commun. ACM **28**(7), 680–689 (1985)
30. Hogan, R., Curphy, G.J., Hogan, J.: What we know about leadership: effectiveness and personality. Am. Psychol. **49**(6), 493 (1994)
31. Hohenstein, J., Jung, M.: AI-supported messaging: an investigation of human-human text conversation with AI support. In: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, pp. 1–6 (2018)
32. Hohenstein, J., Jung, M.: AI as a moral crumple zone: the effects of AI-mediated communication on attribution and trust. Comput. Hum. Behav. **106**, 106190 (2020)

33. Hohenstein, J., et al.: Artificial intelligence in communication impacts language and social relationships. Sci. Rep. **13**(1), 5487 (2023)
34. Jagadeesan, R.: Spontaneous intrapersonal synchrony and the effect of cognitive load. Master's thesis, The University of Western Ontario (Canada) (2023)
35. Jakesch, M., French, M., Ma, X., Hancock, J.T., Naaman, M.: AI-mediated communication: how the perception that profile text was written by AI affects trustworthiness. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–13 (2019)
36. Johansen, R., Vallee, J., Spangler, K.: Electronic Meetings: Technical Alternatives. Addison-Wesley Longman Publishing Co., Inc (1979)
37. Katsh, E., Rifkin, J., Gaitenby, A.: E-commerce, e-disputes, and e-dispute resolution: in the shadow of eBay law. Ohio St. J. on Disp. Resol. **15**, 705 (1999)
38. Kim, A., Cho, M., Ahn, J., Sung, Y.: Effects of gender and relationship type on the response to artificial intelligence. Cyberpsychol. Behav. Soc. Netw. **22**(4), 249–253 (2019)
39. King, J., Just, M.A.: Individual differences in syntactic processing: the role of working memory. J. Mem. Lang. **30**(5), 580–602 (1991)
40. Klapp, O.E.: Overload and Boredom: Essays on the Quality of Life in the Information Society. Greenwood Publishing Group Inc. (1986)
41. Knapp, M.L., Vangelisti, A.L., Caughlin, J.P.: Interpersonal Communication and Human Relationships. Pearson (2014)
42. Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.: lmertest package: tests in linear mixed effects models. J. Stat. Softw. **82**, 1–26 (2017)
43. Langreet, A.: Exploring the impact of workplace friendships on organizational citizenship behavior: the role of social bonds and discretionary effort. Ph. D. thesis (2024)
44. Mair, C.: Using technology for enhancing reflective writing, metacognition and learning. J. Furth. High. Educ. **36**(2), 147–167 (2012)
45. Mairesse, F., Walker, M.: Automatic recognition of personality in conversation. In: Proceedings of the Human Language Technology Conference of the NAACL, Companion Volume: Short Papers, pp. 85–88 (2006)
46. Mairesse, F., Walker, M.A., Mehl, M.R., Moore, R.K.: Using linguistic cues for the automatic recognition of personality in conversation and text. J. Artif. Intell. Res. **30**, 457–500 (2007)
47. Maruping, L.M., Agarwal, R.: Managing team interpersonal processes through technology: a task-technology fit perspective. J. Appl. Psychol. **89**(6), 975 (2004)
48. Mieczkowski, H., Hancock, J.T., Naaman, M., Jung, M., Hohenstein, J.: AI-mediated communication: language use and interpersonal effects in a referential communication task. Proc. ACM Hum. Comput. Interact. **5**(CSCW1), 1–14 (2021)
49. Pennebaker, J.W., Mehl, M.R., Niederhoffer, K.G.: Psychological aspects of natural language use: our words, our selves. Annu. Rev. Psychol. **54**(1), 547–577 (2003)
50. Preacher, K.J., Hayes, A.F.: Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behav. Res. Methods **40**(3), 879–891 (2008)
51. Rahwan, I., et al.: Machine behaviour. Nature **568**(7753), 477–486 (2019)
52. Ravikumar, V., Dhabliya, D., Mathur, M., Das, S., Kumar, R., Rao, S.B.: Ameliorating semantic search through advanced AI techniques. In: 2023 3rd International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), pp. 1–6. IEEE (2023)
53. Reimers, N., Gurevych, I.: Sentence-BERT: sentence embeddings using siamese BERT-networks. arXiv preprint arXiv:1908.10084 (2019)

54. Rule, C.: Online Dispute Resolution for Business: B2B, Ecommerce, Consumer, Employment, Insurance, and Other Commercial Conflicts. John Wiley & Sons (2003)
55. Salimzadeh, S., He, G., Gadiraju, U.: A missing piece in the puzzle: considering the role of task complexity in human-AI decision making. In: Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization, pp. 215–227 (2023)
56. Samson, K., Kostyszyn, P.: Effects of cognitive load on trusting behavior-an experiment using the trust game. PLoS ONE **10**(5), e0127680 (2015)
57. Sprecher, S., Treger, S.: The benefits of turn-taking reciprocal self-disclosure in get-acquainted interactions. Pers. Relat. **22**(3), 460–475 (2015)
58. Sprecher, S., Treger, S., Wondra, J.D., Hilaire, N., Wallpe, K.: Taking turns: reciprocal self-disclosure promotes liking in initial interactions. J. Exp. Soc. Psychol. **49**(5), 860–866 (2013)
59. Stone, P., et al.: Artificial intelligence and life in 2030: the one hundred year study on artificial intelligence. arXiv preprint arXiv:2211.06318 (2022)
60. Thakur, M., Ansari, S., Bidkar, P.: Exploring the relationship between work environment and job satisfaction. Our Heritage **68**(1), 943–956 (2020)
61. Tucker, S., Whittaker, S.: Accessing multimodal meeting data: systems, problems and possibilities. In: International Workshop on Machine Learning for Multimodal Interaction, pp. 1–11. Springer (2004)
62. Vallee, J., et al.: Group communication through computers. Volume 4: social, managerial, and economic issues (1978)
63. Vlahovic, T.A., Roberts, S., Dunbar, R.: Effects of duration and laughter on subjective happiness within different modes of communication. J. Comput.-Mediat. Commun. **17**(4), 436–450 (2012)
64. Weiss, D., Liu, S.X., Mieczkowski, H., Hancock, J.T.: Effects of using artificial intelligence on interpersonal perceptions of job applicants. Cyberpsychol. Behav. Soc. Netw. **25**(3), 163–168 (2022)
65. Yang, Y., et al.: Semantic communications with AI tasks. arXiv preprint arXiv:2109.14170 (2021)
66. Zhang, J., et al.: Conversations gone awry: detecting early signs of conversational failure. arXiv preprint arXiv:1805.05345 (2018)

GlobalUP: Leveraging Social Media to Provide Social Support for Chinese Cultural Groups in Unfamiliar Environments

Haowei Xu(✉) and Xin Shu(✉)

Newcastle University, Newcastle upon Tyne NE1 7RU, UK
{haowei.xu,x.shu6}@newcastle.ac.uk

Abstract. As the largest group of international students, Chinese students play an important role in the UK higher education system. However, little attention has been paid to the difficulties this group encounters in intercultural communication. Both academically and socially, Chinese students are perceived as reticent. Social media has been shown to be an avenue that can be used as social support, but there are some barriers to this approach for Chinese students. For this reason, we explored the possibility of using social media to help build social support in unfamiliar environments and attempted to design an initial prototype social media platform, GlobalUP, to improve cross-cultural communication.

Keywords: Social Media · Social Support · Chinese Cultural Groups

1 Introduction

Globalisation has prompted many Chinese people to migrate overseas for education and work. Chinese students are the largest international student group in the UK, with over 150,000 pursuing degrees in 2023, accounting for one-fifth of the total international student population [11]. While these experiences offer growth prospects, they also present challenges to social integration and mental health due to language barriers and cultural differences [29].

Social support is critical to adjusting to a new environment, promoting physical and mental health, and successful adaptation [20]. Social media can provide this support by facilitating communication and community building [7]. However, existing social media platforms do not account for the gaps created by cultural differences. Communication may be hindered by usage habits and cultural differences [10, 25]. Furthermore, popular Western social media may be less accessible or familiar due to restrictions in China [21], which also ignore Chinese culture when it comes to international students from China, leading to difficulties in use.

This study explores how to provide social support to Chinese communities in unfamiliar environments through a culturally appropriate social media platform. GlobalUP is an initial social media prototype designed to provide social support to Chinese students in unfamiliar environments by building familiarity and trust.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 395–408, 2025.
https://doi.org/10.1007/978-3-031-93539-8_26

2 Related Work

2.1 Cross-Cultural Adaptation Challenges for Chinese International Students

Cross-cultural adaptation is an important phenomenon frequently studied in international student research. U-Curve Model [19] and Culture Shock Theory [22] are well-known theories about the mental stages of changing to a different culture. Kim's Cross-Cultural Adaptation Theory [14] says that adaptation is a moving process. It calls for mental strength and constant learning.

It is particularly difficult for foreign Chinese students to change. This stems from the fact that both their culture and language exhibit large differences. The research pinpointed meaningful difficulties, importantly large language barriers, a large cultural distance, meaningful social isolation and acute mental health issues. Students with limited English skills often battle to fit in socially as well as academically. For help, they may turn to Chinese-speaking friends along with Chinese social media [16,28]. Because social standards and ways of interacting are quite different in Western cultures versus China, Chinese students may have more trouble forming real relationships locally [3,18].

Berry's Acculturation Model [4] categorizes adaptation strategies into the following four types: *integration, assimilation, separation in addition to marginalization.* According to studies, most Chinese students in the U.S. prefer the separation strategy, which means keeping close ties to their own culture and limiting contact with the culture of their host country [16]. This strategy might limit later sociocultural integration, even if it offers some psychological comfort. In addition, a study of Chinese students in the Netherlands emphasized the role of personality traits (e.g., emotional stability, cultural empathy, and social initiative) in predicting successful acculturation [18].

The importance of social support in cross-cultural adaptation has been well documented. Both offline and online social networks play a crucial role in reducing stress and enhancing adaptation [28]. In particular, online ethnic social groups have become an important resource for Chinese students abroad, providing them with emotional and informational support [16,28]. However, Western mainstream social media platforms often fail to meet the cultural needs of Chinese international students, thus contributing to the lack of their digital social experience.

Given these findings, the *GlobalUP* platform is designed to address the cross-cultural adaptation challenges faced by Chinese students by integrating culturally relevant social support features. By leveraging familiar social media interfaces and community-driven interactions, *GlobalUP* aims to bridge the social and psychological gaps in the adaptation process.

2.2 Social Support and Digital Networks for Chinese International Students

Social support plays a crucial role in the psychological well-being and cultural adaptation of international students. It can be broadly categorized into perceived

(subjective) support, which includes emotional and informational resources individuals believe they have, and received (objective) support, which consists of tangible assistance provided by peers, family, and institutions [30]. Studies show that students who feel supported experience less stress, anxiety, and depression, leading to better academic and social adjustment [6,30]. In contrast, a lack of perceived support can result in cultural isolation and reduced life satisfaction [15,23].

Traditionally, social support came from face-to-face interactions with peers, mentors, or local communities [6]. However, in today's digital age, social media has emerged as a key tool for providing social support to international students. Platforms like Facebook, WeChat, and online communities allow students to stay connected with their home culture while forming new social bonds, enhancing their sense of belonging and cross-cultural adaptation [5,17].

Upon further research, it was found that Chinese social media platforms like WeChat and Weibo emphasize closed, relationship-based interactions, where users primarily engage within familiar circles. In contrast, Western platforms such as Facebook and Instagram encourage open communication and networking with strangers. This fundamental difference can lead to adaptation challenges for Chinese students abroad [27]. A study exploring social media usage among Chinese students in the UK found that while these Chinese platforms help maintain cultural ties, they may also limit opportunities for engagement with the host culture [9].

Additionally, privacy concerns shape their online behavior, as Chinese users tend to engage in self-censorship and are more cautious about sharing opinions, especially on Western platforms, where they may feel exposed or worry about misunderstandings due to cultural and linguistic differences [12]. Understanding these cross-cultural differences is essential for designing digital support systems that cater to Chinese students' habits while helping them navigate and integrate into new social environments.

2.3 Culturally Adaptive Design in HCI

Culturally tailored design in Human-Computer Interaction (HCI) recognizes that user experience is deeply influenced by cultural background, shaping interaction preferences, trust, and engagement. Research has shown that adapting user interfaces to align with cultural norms significantly improves usability and user satisfaction [24]. For instance, studies on avatar design found that culturally appropriate representations increase trust and engagement, demonstrating the importance of familiar visual and interaction patterns in digital systems [2]. However, existing social media platforms and social support tools often fail to account for these cultural variations, particularly for Chinese international students. Western platforms like Facebook and Twitter prioritize open, individualistic interactions, whereas Chinese platforms such as WeChat and Weibo emphasize close-knit, relationship-based communication [8]. This fundamental difference can result in adaptation challenges, as Chinese students may struggle to integrate into Western social media ecosystems while maintaining connections

with their cultural identity [13]. Additionally, non-culturally adaptive platforms often fail to support the unique linguistic and social preferences of international students, limiting their ability to seek effective social support [1]. To address these limitations, this study introduces *GlobalUP*, a social media platform that integrates familiar Chinese digital environments with cross-cultural engagement features. By bridging the gap between different interaction paradigms, *GlobalUP* aims to facilitate smoother adaptation and enhanced social support for Chinese students abroad.

3 Prototype Design

The **GlobalUP** prototype was developed as an initial exploration of how a culturally informed social media platform could address the challenges faced by Chinese international students in unfamiliar environments. We designed GlobalUP with three key requirements in mind: (1) *encouraging participation in intercultural activities (R1)*, (2) *creating a familiar environment for information presentation and community (R2)*, and (3) *building user trust and prompting active participation (R3)*.

3.1 Design Rationale and Core Features

Encouraging Participation in Intercultural Activities (R1). To promote active engagement in cross-cultural events, GlobalUP integrates location-based event discovery and an intuitive filtering process (Fig. 1). The *Explore* tab displays real-time event suggestions pinned on a map, emphasizing both Chinese organized and culturally diverse local events (e.g., festivals, workshops). Category filters (e.g., Music, Food & Drink, Workshops) further encourage users to branch out beyond culturally familiar activities. A straightforward booking flow lowers the barrier to joining new events; detailed information (time, location, cost) is provided, and integrated booking keeps users within the app rather than redirecting them elsewhere. These designs aim to reduce anxiety around trying new experiences by making event information transparent and actionable. By prominently featuring a variety of local and cross-cultural events, GlobalUP nudges users to explore beyond their comfort zone.

Creating a Familiar Environment (R2). Many Chinese students rely heavily on WeChat and are accustomed to its interface paradigms, so GlobalUP was designed to align with common Chinese social media usage habits and thereby reduce friction. Users can log in via WeChat, share event information through WeChat groups, and benefit from layouts similar to chat-like social feeds. By adopting recognizable icons and navigational patterns, the learning curve is minimized. We further incorporated a *Friends Event* section, analogous to the "Moments" or "Feed" concept in WeChat, making it easier for new users to locate and share activities within personal networks. By embedding these cultural and

interface cues, GlobalUP supports a sense of familiarity that is critical for international students who may already feel overwhelmed by new social or technical environments.

Building User Trust and Prompting Active Participation (R3). Trust is essential in bridging online and offline interactions. GlobalUP builds trust by displaying user profiles with basic affiliation details (e.g., university name or program) to establish authenticity, and by providing clear privacy controls that let users manage event visibility and participation settings (e.g., public vs. friends-only). Furthermore, user-generated comments and ratings—highlighting, for instance, whether an event is "Chinese friendly" or suitable for first timers offer additional social proof. By reducing social risks and fostering a reliable environment, GlobalUP aims to give Chinese international students the confidence to connect and participate more actively in both culturally familiar and cross-cultural activities.

3.2 Interface Overview

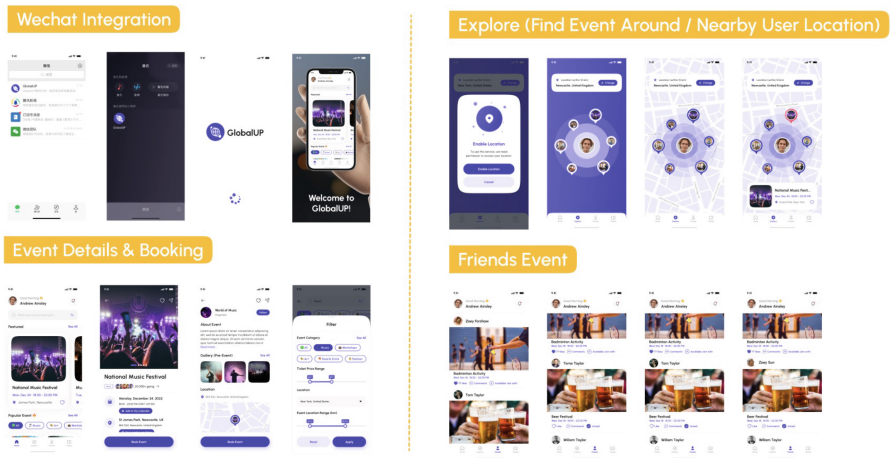


Fig. 1. GlobalUP Prototype Interface.

Figure 1 provides an overview of the GlobalUP interface, illustrating how WeChat integration, event discovery, booking flow, and friend-based recommendations come together:

- **WeChat Integration Screens**: Demonstrate how GlobalUP connects seamlessly with users' existing WeChat accounts and contact lists.
- **Event Details & Booking**: Showcases a cohesive flow from event discovery, details (e.g., images, descriptions, location), to an in-app booking mechanism.

- **Explore**: Depicts location-based functionality, allowing users to see events around their current city or campus. Geolocation data is used to recommend relevant events and highlight friends in proximity.
- **Friends Event**: Emulates a "feed" of events hosted or attended by friends, offering a sense of community and trust by leveraging personal connections.

4 Methods

In order to explore the current social dilemmas faced by Chinese students and validate the usability of the design, we recruited new students from China through advertisements, recruiting newcomers from the past six months and those who have used local social media. Ultimately, we selected ten participants. Most participants had already been living in the UK for three months, with one participant arriving two months early for a pre-semester language course. Participants included five undergraduate and five masters students (Table 1). All participants were fully briefed on the purpose and process of the study prior to its commencement, and participants voluntarily signed an informed consent form. In addition, participants were informed that they could voluntarily withdraw from the study at any time without providing a reason. We ensured that all participants' personal information would be kept strictly confidential and used solely for research purposes. The study was reviewed by the University Ethics Committee.

Table 1. Participant Demographics

ID	Gender	Age	Program	Length of Stay (Months)
P1	Female	19	Computer Science BSc	3
P2	Male	18	Computer Science BSc	3
P3	Male	19	Computer Science BSc	3
P4	Male	18	Computer Science BSc	5
P5	Female	18	Computer Science BSc	3
P6	Male	22	Advanced Computer Science MSc	3
P7	Female	22	Digital Business MSc	3
P8	Male	23	Digital Business MSc	3
P9	Female	23	Digital Business MSc	3
P10	Female	23	Digital Business MSc	3

Our study consisted of three phases; firstly, we focused on exploring each student's perceptions and needs for networking, friendships and social interactions with locals or other international students, as well as their experiences when using local social media. Semi-structured interviews were used and were scheduled to last anywhere from 20 to 40 min per participant, with an average

length of approximately 30 min. The interviews first focussed on the participants' past experiences of engaging in cross-cultural activities. Next, the interviews explored their needs for participation in intercultural activities, with the final part addressing their challenges and the process of acculturation. Afterwards, we used Figma's Interactive prototype and invited the participants to have a 15-minute experience. Our study was not developed using code, but rather followed the Wizard of Oz approach, focusing on user understanding and experience. Afterwards, participants were invited to answer a few short questions, mainly related to their perceptions of social support in the prototype. As compensation for the time and effort invested by the participants, each participant was rewarded with an item equivalent to £10 at the end of the interview. The interviews were conducted in a user research space at the university and all interviews were audio-recorded with the consent of the participants.

After the data had been obtained, the audio data was first converted into a transcript via iFLYTEK, after which the transcript was proofread and corrected. After filtering out invalid data, thematic analysis was performed [26]. Ultimately, the types of theme included problems faced by participants, satisfaction after using prototype, and prospective views on the supportive role of social media.

5 Results

5.1 Compensation: Dilemmas in Communication and Technology Use

Cross-Cultural Limited Opportunities and Barriers to Participation

'I would like to get involved, but I don't know where or how.'

The ten Chinese students who participated in the study generally reported that their interactions with students from other countries or local residents were extremely limited. These interactions mainly took place in classrooms and libraries, and were primarily confined to teaching and learning activities. Consistent with the findings of previous studies, Chinese students did not attend even these activities regularly. In contrast, all participants expressed a desire to engage in more intercultural communication activities. When faced with the opportunity to engage in intercultural activities in the hypothetical scenario, the majority of participants (8 in 10) showed interest and felt that these activities would help to improve their understanding of other cultures, improve their English language skills, and increase their self-confidence, which P8 explained by the fact that interacting with the local community after going abroad should also be part of the study abroad experience, and that it is in interacting with others that one can improve one's language skills.

When asked about the reasons for this phenomenon, some participants noted that it was difficult to find activities. P2 attributed this phenomenon to the fact that their usual contacts were similar to those at home, with the social circle consisting almost exclusively of other Chinese students. This resulted in a social experience similar to life in China, preventing a significant breakthrough

in the cultural barrier and leaving participants unsure of how to connect with locals. P10 described this as a kind of cultural inertia, and a lack of knowledge about how to overcome it. P1 mentioned language barriers, a lack of cultural background knowledge, and the complexity of information about activities. In addition, participants cited safety concerns, especially when attending activities alone, fear of safety risks and potentially embarrassing situations, which further affected their willingness to engage in activities.

Social Media Use Dilemma

'I tried, but I felt like I was playing a single-player game.'

Differences in social media use significantly worsened participants' limitations in cross-cultural socialisation. With the exception of P4 and P5, most participants did not attempt to find cross-cultural socialising activities through local platforms (e.g., Facebook), instead relying primarily on Chinese platforms such as WeChat and Xiaohongshu for information about activities. This reliance on specific social platforms limited their social opportunities and made it difficult to access activities related to local culture or other international student groups. Since most of these platforms serve the Chinese community, participants' access to information about activities and interactions were mainly centred around Chinese student circles, thus limiting their opportunities to integrate into multicultural environments. P7 further explained that the activities found in these communities were mostly attended by Chinese people. In contrast, participants failed to develop the habit of using Western social media such as Facebook and Instagram. In this regard, P2 and P10 attributed the discomfort to poor user experience, and P4 argued that the lack of guidance and support from local friends made participants less interested and motivated to make connections on new social platforms. As a result, most participants' social activities were centered around interactions with other Chinese students.

5.2 Satisfaction: Increased Trust and Familiarity

Familiarity Promotes Acceptance and Use

'The familiarity of the software makes me want to use it more'

All participants were able to successfully master the operation of the prototype when using it for the first time and were able to use GlobalUP accurately, and the majority of the participants showed a high level of autonomy and interest in exploring the various features of the prototype. Only P4 relied more on external guidance during operation, while the others demonstrated strong independent learning abilities. P2 and P3 specifically mentioned that integrating WeChat's inbuilt functionality into the prototype design lowered the learning curve, enabling them to adapt quickly and use the prototype effectively. The familiarity element in the prototype design, especially the similarity to the functionality of common social platforms such as WeChat, may have played an important role in technology acceptance. P9 explained that the adoption of a social

platform design similar to WeChat not only increased the frequency of her use of the prototype, but also increased her willingness to participate in the activities.

Sense of Trust Encourages Participation

'I trust because it's familiar to me.'

The majority of participants showed interest in the cross-cultural socialising activities demonstrated in the prototype and generally expressed a high level of trust in the platform. Participants felt that the prototype was able to provide effective social support and facilitate social interactions, showing the potential of the prototype in these areas. P4 specifically mentioned that the way the information was presented in the prototype enhanced their trust in the platform, mainly because it was presented in a way that was similar to the social platforms that they use on a day-to-day basis, which lowered the psychological burden of the process of using it. Through this feedback, it can be surmised that the sense of intimacy in the platform's design may enhance users' trust through the similarity of the interface layout and functional design. This design may reduce the uneasiness and discomfort that users may feel when adapting to a new platform, thus increasing the acceptability of the platform. In addition, as familiar design elements enable users to adapt to the prototype in a shorter period of time, this may make them feel more at ease and comfortable in the process of using it, which in turn enhances their trust and acceptance of the platform. Thus, the intimacy of the platform's design may have played a key role in increasing user trust and lowering barriers to use.

Participants'feedback on the way information was presented in the prototype suggested that the form in which the information was presented had an impact on users' motivation to participate. P1 tended to prefer presentations combining text and images, believing that this form was more effective in capturing users' attention and boosting their motivation to participate. P3 specifically mentioned that the exact time and location of the event was crucial in engaging participants and could increase the event's transparency and thus increase participants' trust. The clear presentation of information can reduce the cognitive load of users when participating in an event and help them better understand the programme. This suggests that accurate and clear presentation of information plays an important role in the attractiveness and transparency of an event and may directly influence users' participation decisions. In addition, P1 showed interest in the 'Show friends to join the event' feature of the prototype, believing that joining the event with friends would help increase their motivation to participate and the social support they receive. This feedback suggests that the design of social features may increase participants' willingness to participate by facilitating their interactions with friends or others. However, P1 also noted that he would feel his privacy was threatened if the activity was open to everyone, highlighting the importance of privacy protection in the design of social platforms. Therefore, the way information is presented in the platform, the design of social features, and privacy protection measures may jointly affect users' willingness to participate. By increasing the transparency of activities, enhancing social interactions,

and ensuring the security of user privacy, it is possible to promote more active participation.

5.3 Transcend: Reconstruction of Self-identity

In this study, the GlobalUP prototype demonstrates the potential of social media in helping Chinese international students to cope with the challenges of cross-cultural communication, and also provides self-identity. P4 explains that through GlobalUP, international students have the potential to be able to discover and engage in activities with local friends, thus reducing the sense of isolation that comes with living in a foreign country. The social features in the platform enable students to connect with groups of people with the same cultural background, which may be able to help them maintain an emotional connection. The sense of achievement that comes from participating in activities goes beyond the connection that the design brings and can lead to personal identity, which P4 explains as a remarkable feeling of being in an unfamiliar environment, and a sense that one has finally done it, and that one has finally studied abroad, rather than in another homeland.

6 Discussion

In this section, we discuss how our findings contribute to existing cross-cultural adaptation theories and HCI practices, as well as the implications, limitations, and directions for future work.

6.1 Contributions to Cross-Cultural Adaptation Theories

Our study corroborates and extends key insights from prominent cross-cultural adaptation frameworks, including Berry's Acculturation Model [4] and Kim's Cross-Cultural Adaptation Theory [14]. Specifically, the results show that when Chinese international students are provided with a social media environment that incorporates culturally familiar interaction patterns, they are more willing to move beyond separation oriented behaviors (e.g., only socializing with co-nationals) and progress toward integrative strategies.

In line with Kim's emphasis on the interplay between stress and adaptation, GlobalUP's familiar interface appears to reduce the stress of using an unfamiliar platform, thus easing the initial phase of psychological adaptation. Additionally, our findings resonate with Berry's view that individuals are more likely to adopt an integration strategy when they retain access to familiar cultural resources yet have pathways to engage with the new culture. By presenting event information in a culturally tailored and transparent manner, GlobalUP effectively lowers the barrier to local participation, thereby helping users expand their social circles and enhance cross-cultural interactions.

6.2 Implications for HCI and Social Media Design

Culturally Adaptive Interfaces. Our work underlines the importance of designing interfaces that leverage existing user habits and cultural preferences. The GlobalUP prototype adopts recognizable elements from Chinese social media platforms such as WeChat (e.g., interface layouts, messaging functionalities), which quickly established user trust and engagement. Future HCI research could further refine this design approach by systematically identifying and integrating region specific cultural cues into social platforms targeting international populations.

Balancing Familiarity and Exploration. While familiarity facilitates user adoption and reduces the anxiety of engaging with new tools, it is equally critical to encourage exploration beyond one's cultural comfort zone. GlobalUP provides culturally familiar communication channels alongside features (e.g., curated local events, friend activity visibility) that nudge users to connect with non-Chinese peers. This interplay between "comfort" and "exploration" can be generalized to other cross-cultural designs where the goal is to gently guide users toward broader social integration rather than abrupt immersion.

Supporting Activity Discovery and Trust. In our user interviews, participants consistently noted the challenge of discovering credible multicultural activities and the importance of transparency in event details. By providing explicit information such as event descriptions, schedules, and friend endorsements, GlobalUP builds a sense of trust, thereby increasing the likelihood of participation. From an HCI standpoint, incorporating social proof (e.g., "friends who joined") and verifying event authenticity can play a critical role in online community-building, particularly for populations new to a cultural context.

6.3 Limitations and Future Work

Despite these contributions, our research has several limitations. First, the sample size (ten participants) was relatively small and recruited primarily from a single university setting; therefore, the results may not generalize to other cultural or academic contexts. Future studies should include a larger and more diverse participant pool potentially across different countries or academic programs to validate and refine the design features identified here.

Second, our Wizard-of-Oz style prototype provided a controlled environment that may not fully capture the complexities of long term usage. Follow up longitudinal field studies with a fully implemented GlobalUP platform could offer deeper insights into how students' behaviors, engagement levels, and cross-cultural adaptation evolve over time. Additionally, real-world deployment would allow for more robust examination of privacy concerns, data security, and potential trade offs between localized cultural features and mainstream social media integration.

Lastly, we focused specifically on Chinese international students. While the culturally adaptive design principles could be transferred to other groups, further research is needed to identify which design elements (e.g., language support, interface layouts, cultural cues) are most effective for other international or minority populations. Comparative studies across different cultural groups would further enrich our understanding of how to balance culturally specific features and universal usability principles in global HCI designs.

7 Conclusion

This study highlights the importance of culturally tailored social media platforms in providing social support for Chinese communities in unfamiliar environments. Our findings illustrate how GlobalUP's WeChat inspired interface and culturally relevant content can significantly reduce adoption barriers, foster trust, and encourage cross-cultural participation by offering transparent event details, social proof, and friend endorsements, thereby addressing key user concerns around activity discovery and safety. Going forward, we plan to refine the GlobalUP prototype into a fully functional system, explore automated event recommendation mechanisms, and conduct larger scale deployments to investigate real world usage patterns, which will not only test the platform's scalability and long-term impact but also reveal how such interventions can systematically support international students' well-being and integration across diverse cultural contexts. By continuing to build upon GlobalUP's culturally adaptive design approach and addressing the challenges identified in this work, we envision the platform evolving into a valuable resource for Chinese users and potentially other cultural groups as they adapt to new social landscapes.

References

1. Joo, J.Y., Liu, M.F.: Culturally tailored interventions for ethnic minorities: a scoping review. Nurs. Open **8**(5), 2078–2090 (2020). https://doi.org/10.1002/nop2.733
2. Aljaroodi, H.M., Adam, M.T.P., Teubner, T., Chiong, R.: Understanding the importance of cultural appropriateness for user interface design: an avatar study. ACM Trans. Comput. Hum. Interact. **29**(6), 52 (2022)
3. Anonymous: Factors affecting Chinese short-term international students' cross-cultural adaptation in psychology, learning, and life. Int. J. Intercultural Relat. **50**, 97–113 (2022)
4. Berry, J.W.: Immigration, acculturation, and adaptation. Appl. Psychol. **46**(1), 5–34 (1997)
5. Cheng, Y., Meng, J., Liu, S.: Personal network structure and perceived social support in the context of intercultural adjustment. Commun. Q. **66**(4), 576–594 (2018)
6. Chui, R.C., Chan, C.K.: School adjustment, social support, and mental health of mainland Chinese college students in Hong Kong. J. Coll. Stud. Dev. **58**(1), 100–88 (2017)

7. Dumaru, P., et al.: Understanding the transitional challenges with technology use after relocating to the USA. In: International Conference on Human-Computer Interaction, pp. 201–220 (2023)
8. Gevorgyan, G., Manucharova, N.: Does culturally adapted online communication work? a study of American and Chinese internet users' attitudes and preferences toward culturally customized web design elements. J. Comput.-Mediat. Commun. **14**(2), 393–413 (2009)
9. Guo, W.: Navigating cultural integration: the role of social media among Chinese students in the UK. J. Knowl. Econ. 1–32 (2024)
10. Han, K., Volkova, S., Corley, C.D.: Understanding roles of social media in academic engagement and satisfaction for graduate students. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, pp. 1215–1221 (2016)
11. Higher Education Statistics Agency (HESA): Where do students come from? (2024). https://www.hesa.ac.uk/data-and-analysis/students/where-from
12. Hu, Y., Chen, H.: Silence on social media: factors shaping Chinese international students' decision on speaking up on social media. J. Intercult. Commun. Res. **52**, 1–25 (2022)
13. Huang, K.H., Deng, Y.S.: Social interaction design in cultural context: a case study of a traditional social activity. Int. J. Des. **2**(2), 81–96 (2008)
14. Kim, Y.Y.: Becoming Intercultural: An Integrative Theory of Communication and Cross-cultural Adaptation. Sage Publications (2001)
15. Kristiana, I.F., et al.: Social support and acculturative stress of international students. Int. J. Intercultural Relat. (2022)
16. Lai, H., Wang, D., Ou, X.: Cross-cultural adaptation of Chinese students in the united states: acculturation strategies, sociocultural, psychological, and academic adaptation. Front. Psychol. **13**, 924561 (2023)
17. Lin, J., Peng, W., Kim, M., Kim, S.Y., LaRose, R.: Social networking and adjustments among international students. New Media Soc. **14**(3), 421–440 (2012)
18. Long, J.H., Yan, W.H., Yang, H.D., van Oudenhoven, J.P.: Cross-cultural adaptation of Chinese students in the Netherlands. US-China Educ. Rev. **6**(9), 58–72 (2009)
19. Lysgaard, S.: Adjustment in foreign society: Norwegian fulbright grantees visiting the United States. Int. Soc. Sci. Bull. **7**, 45–51 (1955)
20. Morioka, T.T., Boland, J.: Identity construction and social capital: a qualitative study of international students' social networks in Japan. Int. J. Intercult. Relat. **55**, 1–11 (2016)
21. Motahar, T., et al.: Exploring how international graduate students in the us seek support. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1–8 (2020)
22. Oberg, K.: Cultural shock: adjustment to new cultural environments. Pract. Anthropol. **7**, 177–182 (1960)
23. Razgulin, J., et al.: The role of social support and sociocultural adjustment for international students' mental health. Sci. Rep. (2023)
24. Reinecke, K., Bernstein, A.: Improving performance, perceived usability, and aesthetics with culturally adaptive user interfaces. ACM Trans. Comput. Hum. Interact. **18**(2), 1–33 (2011)
25. Skaržauskienė, A., Tamošiūnaitė, R., Žalėnienė, I.: Defining social technologies: evaluation of social collaboration tools and technologies. Electron. J. Inf. Syst. Eval. **16**(3), 231–240 (2013)

26. Tenny, S., Brannan, J.M., Brannan, G.D.: Qualitative study (2017)
27. Xu, Q., Mocarski, R.: A cross-cultural comparison of domestic American and international Chinese students' social media usage. J. Int. Students (2014)
28. Ye, J.: Traditional and online support networks in the cross-cultural adaptation of Chinese international students in the United States. J. Comput. Mediat. Commun. **11**(3), 863–876 (2006)
29. Yu, B., Wright, E.: Socio-cultural adaptation, academic adaptation and satisfaction of international higher degree research students in Australia. J. Int. Stud. **6**(3), 1–13 (2016)
30. Yusoff, Y.: Self-efficacy, perceived social support, and psychological adjustment in international undergraduate students in a public higher education institution in Malaysia. J. Stud. Int. Educ. **16**(4), 353–371 (2012)

Author Index

A
Aguilera, Ricardo II-95
Alimerko, Rudina I-115, I-141
Ancán-Bastías, Oscar I-3
Ando, Masayuki I-360
Aroca, Geraldine I-272
Astudillo, Agustín I. I-37
Azocar, Isidora I-272

B
Badejo, Johannes A. I-303
Baldauf, Matthias II-31
Bashir, Asim I-125
Bazlutckaia, Mariia II-189
Bian, Andrew II-343
Björnsdóttir, Kristín I-194
Bodrunova, Svetlana S. I-351, II-210
Botella, Federico I-219
Brooks, Eva Irene I-194

C
Cano, Sandra I-236
Cares, Carlos I-3
Castillo, Tomás I-20
Castro, John W. I-37, I-87
Catalán, Ignacio I-20
Chakraborty, Joyram I-303
Chamorro, Benjamín I-20
Chen, Yiwei II-3
Chuang, Miao-Hsien II-21
Collis, Avon I-87
Cornide-Reyes, Héctor II-228
Çupi, Deniz I-141

D
Dau, Susanne I-194
Debrah, Ralitsa I-257
Díaz-Arancibia, Jaime I-3
Durán, Eduardo II-228

E
Ezzat, Mahinour I-172

F
Farouqa, Georgina I-155, I-180
Federspiel, Esther II-31
Fedorov, Artem II-304
Fedyanin, Denis II-247
Feng, Xi II-72

G
Godoy, María Paz O. I-272
Gross, Tom I-335
Guo, Qihan II-48

H
Háhn, Judit I-194
Hamam, Doaa I-172
Hashemi, Sylvana Sofkova I-194
Hatamleh, Zaydoon I-155
Hiasat, Laith I-155
Hysaj, Ajrina I-125, I-155, I-172, I-180

I
Ibrahim, Othman I-155
Ibrayeva, Galiya II-257
Inui, Ryotaro II-272
Ischenko, Roman II-304
Ishikawa, Aina I-57
Isradisaikul, Chinawat I-290
Izumi, Tomoko I-360

J
Jeong, Hongseok II-343
Jóhannesdóttir, Sigurbjörg I-194
Jugo, Gordana I-319

K
Kamegne, Yvonne II-62
Kasymov, Renat II-293

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Switzerland AG 2025

A. Coman and S. Vasilache (Eds.): HCII 2025, LNCS 15786, pp. 409–411, 2025.
https://doi.org/10.1007/978-3-031-93539-8

Khan, Sara Azeem I-155, I-172, I-180
Kimmerle, Joachim II-83
Kjartansdóttir, Skúlína Hlíf I-194
Knieling, Victoria II-343
Kongsawasdi, Siriphan I-290
Korchagina, Anna II-304
Koren, Leon I-70, I-319, II-177
Korkala, Salme I-194
Kousa, Päivi I-194
Kragic, Danica I-97
Kuttal, Sandeep Kaur II-322

L
Lee, Kyoungyong II-48
Leon, Shahnewaz II-322
Li, He II-48
Li, RuiChi II-48
Lindström, Nataliya Berbyuk I-194
Liu, Lili II-72
Lundin, Mona I-194
Luo, Fei I-20, I-87
Luther, Teresa II-83

M
Maleck, Moritz I-335
Marin, Vicente I-87
Matsuo, Naoki II-282
Matus, Nicolás I-219, II-95, II-109
Matus, Sebastian I-236
Maysuradze, Archil II-304
McKinney, Brett II-322
Meng, Qingshen II-48
Monsalves, Diego II-228
Morales, Jenny II-228

N
Namatame, Takashi II-123, II-272, II-282, II-360
Nigmatullina, Kamilla II-293
Nomaan, Saima I-125

O
Oladapo, Helina I-257, II-62
Orsag, Luka I-70, II-177
Orthmann, Bastian I-97
Ortiz, Oriella II-95, II-109
Otake, Kohei I-57
Otsu, Kouyou I-360
Ouda, Yasser I-172
Owais, Amjad I-155
Owusu, Eric I-257, II-62, II-162
Oyedeji, Shola I-194

P
Peansukmanee, Siriporn I-290
Peko, Ivan I-319
Peral Cortés, Jesús I-3
Peter, Andreas II-31
Piñones, María Angélica S. I-272
Pratama, I. Gede Putu Dinanda Putra II-21
Punchenko, Petr II-189

R
Ragavan, Sruti Srinivasa II-322
Ramazan, Aigul II-257
Rhodes, Geoffrey Alan II-3
Rink, Olga II-304
Rodicheva, Alena I-351
Rojas, Luis I-20, I-37, I-87
Rusu, Cristian I-219, II-95, II-109
Rusu, Virginica II-95, II-109

S
Saavedra, Camila II-95
Schaffner, Dorothea II-31
Schwarzenberg, Pablo I-20, I-87
Sedhain, Abim II-322
Shao, Mingchen II-343
Shi, Mingxi II-48
Shu, Xin I-395
Silva-Aravena, Fabián II-228
Sjöberg, Jeanette I-194
Sommer, Michel II-83
Sorbello, Katrina I-87
Stipancic, Tomislav I-70, I-319, II-177
Støckert, Robin I-194
Su, Vanessa II-343
Sytnik, Anna II-189

T
Tabachenkov, Andrey II-304
Takanami, Joshujio I-57
Takata, Nonoka II-123
Tapia, Sergio II-95
Tarvainen, Jenny I-194
Thakur, Nirmalya II-343
Thovuttikul, Sutasinee I-290
Tian, Yuanna II-72

Tita, Miselta W. I-303
Tleugazina, Diana II-257
Tsvetkov, Timofey II-189
Tsybanov, Ilya II-304

V
Van Haastregt, Jonne I-97
Vorontsov, Konstantin II-304

W
Wang, Panpan II-72
Wantanajittikul, Kittichai I-290
Watababe, Tomoro I-360
Welle, Michael C. I-97
Wen, Ouyang II-135
Wu, Jingying II-72

X
Xhoga, Deniz Çupi I-115
Xin, Hu II-135
Xu, Haowei I-395
Xu, Xinyue I-375

Y
Yang, Jianhua II-3
Yeboah, Nana II-162
Yoshimi, Tomoki II-360

Z
Zhang, Peng II-48
Zhang, Yuchong I-97
Zheng, Ruonan I-375
Zhong, Yiting II-72
Zidaric, Matija I-70, II-177