Published in Journals: Education Sciences, Economies, Sustainability and Applied System Innovation

Topic Reprint

Effectiveness and Sustainable Application on Educational Technology

Edited by
Jian-Hong Ye, Yung-Wei Hao, Yu-Feng Wu and Savvas A. Chatzichristofis

mdpi.com/topics

MDPI

Effectiveness and Sustainable Application on Educational Technology

Effectiveness and Sustainable Application on Educational Technology

Editors

Jian-Hong Ye
Yung-Wei Hao
Yu-Feng Wu
Savvas A. Chatzichristofis

MDPI

Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Editors
Jian-Hong Ye
Faculty of Education
Beijing Normal University
Beijing
China

Yung-Wei Hao
Graduate Institute of Curriculum and Instruction
National Taiwan Normal University
Taipei City
Taiwan

Yu-Feng Wu
Office of Physical Education
Ming Chi University of Technology
New Taipei City
Taiwan

Savvas A. Chatzichristofis
Department of Computer Science
Neapolis University Pafos
Pafos
Cyprus

Editorial Office
MDPI AG
Grosspeteranlage 5
4052 Basel, Switzerland

This is a reprint of articles from the Topic published online in the open access journals *Sustainability* (ISSN 2071-1050), *Education Sciences* (ISSN 2227-7102), *Applied System Innovation* (ISSN 2571-5577), and *Economies* (ISSN 2227-7099) (available at: https://www.mdpi.com/topics/sustainable_education).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-7258-2465-6 (Hbk)
ISBN 978-3-7258-2466-3 (PDF)
doi.org/10.3390/books978-3-7258-2466-3

Contents

About the Editors

Jian-Hong Ye

Dr. Jian-Hong Ye is currently a lecturer in the Faculty of Education at Beijing Normal University, a distinguished expert at the National Institute of Vocational Education at Beijing Normal University, and a visiting professor at Ningxia Polytechnic, China. He graduated from the Department of Industrial Education at National Taiwan Normal University, Taiwan, where he also served as a postdoctoral researcher. Dr. Ye is an honorary member of the Phi Tau Phi Scholastic Honor Society (Taiwan) and an outstanding alumnus of the Department of Fashion Design at Tainan University of Technology, Taiwan. In recent years, Dr. Ye has served as an associate editor for a section of Heliyon, and he has also worked as a guest editor and peer reviewer for several SSCI, SCI, and Scopus journals. He enjoys conducting empirical research from a psychological perspective. His research interests include the digitalization of vocational education, curriculum and teaching innovation, student psychological health and well-being, and interdisciplinary topics.

Yung-Wei Hao

Dr. Yung-Wei Hao is a professor at the Institute of Curriculum and Instruction at National Taiwan Normal University in Taipei, Taiwan. Professor Hao is also an instructional designer specialized in online learning environments and the evaluation of instructional technology products and projects. Professor Hao's main research interests include teaching English as a foreign language, inquiry-based learning, game-based learning, effective online instructional strategies, interactivity in online learning, web-based training, and foreign language education through online learning environments.

Yu-Feng Wu

Dr. Yu-Feng Wu is a faculty member at Ming Chi University of Technology, Taiwan, serving as an Assistant Professor in the Office of Physical Education. Driven by a passion for sports management, Dr. Wu earned his doctoral degree from the Graduate Institute of Sport, Leisure, and Hospitality Management at National Taiwan Normal University (NTNU), Taiwan. After graduation, he served as a postdoctoral researcher at the Physical Education Research and Development Center at NTNU. Dr. Wu's research interests include sport psychology, sports marketing, consumer behavior, physical education, digital learning, exergaming, and contemporary hot and emerging educational issues. With this broad spectrum of research interests, he has served as a guest editor and peer reviewer for several Social Science Citation Index (SSCI), Science Citation Index (SCI), and Scopus journals. Beyond his research pursuits, Dr. Wu is dedicated to excellence in teaching and embodies a commitment to advancing knowledge and fostering intellectual curiosity within the academic community at Ming Chi University of Technology.

Savvas A. Chatzichristofis

Dr. Savvas A. Chatzichristofis is the Vice-Rector for Research and Innovation and a Professor of Artificial Intelligence at Neapolis University Pafos, Cyprus, where he also directs the Intelligent Systems Laboratory. He holds a Ph.D. in Computer Vision and has extensive experience in AI, computer vision, and robotics. His research spans several high-impact projects, including pioneering autonomous navigation systems for UAVs and contributions to image retrieval technologies. He has been recognized with numerous awards and is actively involved in educational robotics, where his work has garnered multiple distinctions. A senior IEEE member, Prof. Chatzichristofis also serves on various editorial boards and contributes to international conferences.

Preface

With the rapid development of educational technology over the past two decades, it has become a focal point for governments, academia, educational experts, teachers, and parents worldwide, thereby driving a wave of research in this field and further highlighting the central role of educational technology in modern education. In-depth exploration of research in this field will uncover gaps in the theory and practice of contemporary educational technology and provide key insights. Therefore, the theme "Effectiveness and Sustainable Application of Educational Technology" covers multiple thought-provoking research topics, including the online self-learning motivation of Chinese university students during the pandemic, an evaluation method for sustainable enrollment planning in Chinese universities based on Bayesian networks, the impact of blended teaching strategies on the performance of first-year English major students, and a literature review on generative artificial intelligence in educational settings. These studies expand our understanding of the application of educational technology. The research findings related to this theme will help address questions regarding the theoretical and practical applications of contemporary educational technology. It is hoped that after reading this reprint, readers will further explore how to effectively apply technology in different educational/learning settings to facilitate the cultivation of knowledge and skills from various theoretical perspectives in the field of educational technology.

Funding: This work was supported by the Fundamental Research Funds for the Central Universities in China (grant number: 2022NTSS52), First-Class Education Discipline Development of Beijing Normal University (grant numbers: YLXKPY-XSDW202408, YLXKPY-XSDW202211, YLXKPY-ZYSB202201) and 2024 Beijing Normal University's Teachers' Teaching Development Fund Project (grant number: 2024125).

Jian-Hong Ye, Yung-Wei Hao, Yu-Feng Wu, and Savvas A. Chatzichristofis
Editors

sustainability

Editorial

Effectiveness and Sustainable Applications of Educational Technology

Jian-Hong Ye [1,2,†], Yung-Wei Hao [3,†] and Yu-Feng Wu [4,*]

1 Faculty of Education, Beijing Normal University, Beijing 100875, China; jianhong.ye@bnu.edu.cn or kimpo30107@hotmail.com
2 National Institute of Vocational Education, Beijing Normal University, Beijing 100875, China
3 Graduate Institute of Curriculum and Instruction, National Taiwan Normal University, Taipei City 106, Taiwan; hao@ntnu.edu.tw
4 Office of Physical Education, Ming Chi University of Technology, New Taipei City 24301, Taiwan
* Correspondence: garywu821@mail.mcut.edu.tw
† These authors contributed equally to this work and share first authorship.

Citation: Ye, J.-H.; Hao, Y.-W.; Wu, Y.-F. Effectiveness and Sustainable Applications of Educational Technology. *Sustainability* **2024**, *16*, 8209. https://doi.org/10.3390/su16188209

Received: 5 September 2024
Accepted: 12 September 2024
Published: 20 September 2024

Over the past few decades, the development and application of educational technology has drawn the attention of governments, academic organizations, educators, teachers, and parents around the world. With the advent of the Industry 4.0 era, the digital transformation of education has also been put on the agenda, making educational technology a popular topic in educational research. Due to the continuous innovation and development of educational technology, as well as the ongoing emergence and evolution of educational technology theories, coupled with the significant impact of the COVID-19 outbreak in early 2020 on education, there has been a widespread realization of the importance of issues related to educational technology.

Today, digital technology has become an important tool for achieving the goal of providing inclusive and equitable quality education [1]. Educational technology is considered to provide learners with opportunities to receive high-quality education. If used properly, these technologies can enable learners to study anytime and anywhere in a virtual environment, unrestricted by time and space, creating meaningful learning experiences for them. From the above, it is evident that new technologies in the field of education are transforming our ideas, perceptions, and educational goals, not just our methods [2]. With the increase in research on the use of educational technology, there is growing attention to the psychological processes involved in teaching and learning with these tools [3]. Since the success of educational technology is influenced by the features of human cognitive style, it is possible for the effectiveness of technology-based education used without reference to instructional design ideologies derived from human cognition to be random [4]. Therefore, to empower students for holistic development through technology, it is essential to understand how teachers can effectively apply technology-assisted teaching and how students can effectively utilize technology-assisted learning. This includes educational data mining and analysis, evaluation and feedback on teaching effectiveness, suggestions for learning progress and methods, the comprehensive recording of multidimensional learning journeys, multimodal learning/learning situation analysis, and providing personalized learning plans based on individual circumstances.

The modern world requires more efficient learning models that enable students to play a more active role in their education [5]. Moreover, these learning approaches need to be oriented towards competency development. Digital literacy encompasses a range of skills, knowledge, abilities, and attitudes related to digital technologies. Digital literacy is considered one of the core competencies essential for the 21st century, contributing to holistic development and lifelong learning. Therefore, how to effectively cultivate digital literacy among both teachers and students has become a key issue. Additionally, topics such as smart education, programming education, competency-based education, interdisciplinary education, online learning communities, blended learning communities, digital (evaluation) literacy of teachers, cloud-based textbooks, digital textbooks, teaching resource libraries, digital platforms, learning societies, lifelong learning, lifelong education, distance education, credit banks, multimodal analysis, educational data mining, big data analysis models, data evaluation, the misuse of technology, technology dependency, ineffective use of technology, long-term impacts of technology, technology transparency, user experience, core competency education, digital literacy, media literacy, and digital competence are also highly prominent research subjects in contemporary education.

With the rapid development of metaverse technology, educators urgently need to understand how to apply learning theories to technology-driven learning experience design. The design of effective metaverse technology learning experiences largely relies on a deep understanding and integration of learning theories. Incorporating educational theories into the design process aims to create meaningful and engaging learning experiences to attract students and promote more effective learning outcomes. The use or implementation of educational technology is not the sole determinant of educational outcomes. Factors such as curriculum design, teaching methods, student characteristics, tool effectiveness, and other variables also influence the educational experience. Thus, research topics on immersive learning, interactive learning, self-directed learning, personalized learning, collaborative learning, project-based learning, problem-based learning, mobile seamless learning, game-based learning, blended learning, inquiry-based learning, ubiquitous learning, ideological and political-based learning (IPL), information technology and interdisciplinary learning (i-STEAM), flipped classrooms, and other student-centered learning/teaching approaches, or online teaching/learning theory models, are receiving significant attention from educators and field scholars. They continuously explore effective practices, which in turn fosters ongoing innovation and development in digital learning theories.

In the theme of "Effectiveness and Sustainable Application on Educational Technology", there are numerous research topics worth referencing. These include online self-learning motivation among Chinese college students during the pandemic, nonlinear teaching methods in economics, self-assessment of teachers' use of mobile devices in the context of gender awareness, evaluation methods for sustainable enrollment planning in Chinese universities based on Bayesian networks, the impact of blended teaching strategies on first-year English majors' performances, a literature review on generative artificial intelligence in educational environments, the role of dynamic geometry software in teacher–student interactions, the meta-analysis of the effectiveness of educational robots in improving learning outcomes, challenges and opportunities in applying identification technology to biology courses, factors influencing Jordanian scholars' behavioral intentions to use sustainable cloud-based quality management systems, understanding students' views on sustainability using natural language processing techniques, a literature review on adaptive learning systems, and the impact of electronic human resource management. These studies expand our understanding of new applications in educational technology. Research findings related to this theme will help address questions about both the theoretical and practical applications of contemporary educational technology.

In the new era of rapid artificial intelligence development, future talent will need to possess interdisciplinary knowledge and skills, enabling them to innovate alongside AI and build diverse professional expertise. However, in recent months, there has been increasing concern in academia about the use of generative artificial intelligence (AI), with one of the main worries being that students may use generative AI tools to cheat or plagiarize their written assignments and exams [6]. Improper use of Artificial Intelligence Generative Content (AIGC) can lead to a lack of critical thinking, plagiarism, and complete reliance on technology by users. Therefore, in the era of artificial intelligence, it is important to consider how people can avoid being replaced by AI and robots and focus on developing high-quality, highly skilled talent. Education should increasingly aim to cultivate students' soft skills, such as innovation and critical thinking.

Technologies such as artificial intelligence and the metaverse have fundamentally changed the thinking, behavior, and lifestyles of stakeholders in the educational environment. This represents an irreversible trend. AIGC can generate human-like responses to interact with users, including answering questions, providing feedback, and guiding learners through their tasks. For teachers, AIGC tools can replace them in performing mechanical and repetitive tasks, such as quickly generating teaching resources, analyzing student progress, assisting with classroom discussions, and participating in assignment evaluations. This allows teachers to instinctively free themselves from routine teaching duties.

At the same time, using AIGC to create videos for visualizing instructional content represents a breakthrough in technology. It enables the possibility of virtual hosts acting as teachers, allowing people to exist in a virtual world. At the same time, it is crucial to pay special attention to individuals' personal rights, such as their likeness and voice. This underscores the importance of technology and digital ethics. Additionally, it is important to address how to effectively disseminate knowledge through online media while safeguarding privacy and intellectual property. This includes efficiently applying technology for knowledge sharing, cultural transmission and preservation, scientific knowledge popularization, and ensuring high-quality knowledge sharing, diffusion, transformation, innovation, and documentation. Strict respect for and emphasis on technological ethics are essential in managing educational data and privacy protection.

Research on individual differences among students helps understand the learning difficulties they might encounter during the learning process. This understanding can improve teaching methods, provide students with better learning advice, offer more suitable teaching content, and enhance teaching approaches [7]. Moreover, how to design data-driven teaching activities based on generative artificial intelligence to meet the needs of improving emotional experiences and promoting cognitive learning in online education has not yet been thoroughly discussed. Furthermore, not everyone is able to adapt to and use new technology effectively. Issues such as user behavior, user engagement, effectiveness of use, users' values regarding educational technology products, and their understanding of educational technology are important topics that deserve in-depth investigation. At the same time, user experience theories and related variable validations for metaverse technologies, including AR/VR/MR/SR/XR and AI, such as AI acceptance models, AI usage effectiveness, AI usage attitudes, AI usage expectations, AI experience value, AI trust, AI illusion, AI dependence, AI effectiveness, AI-related concerns, AI usage intentions, AI literacy, AI-TAPCK, AI learning partners, and other topics, also require ongoing exploration to address educational application issues in the era of technological advancement. Additionally, UNESCO released the "*Guidance for generative AI in education and research*" [8], which examines the significance of generative AI in education from aspects such as governance rules, educational regulation and application policies, and the design and evaluation of educational practices.

In the current educational field, data-driven educational management and digital leadership are increasingly becoming key factors in improving educational quality. Beyond traditional teaching and learning, it is also crucial for educational administrators to effectively use technology for management and for educational supervisors to leverage technology for evaluation. The effective application of technology can optimize the allocation of educational resources, enhance management efficiency, and provide more precise data support for evaluations, thereby advancing the overall development of the education system. Through in-depth data analysis and intelligent management tools, the education system can better address challenges, implement more scientific decision-making, and ultimately achieve comprehensive improvements in educational quality. Therefore, topics such as data-driven educational management and digital leadership are also among the current hot issues.

As we face the arrival of the AI era, education must play a key role in helping people develop global awareness, humanitarian concerns, and a sense of well-being. Technological innovations and transformations continuously reshape human society and civilization, further promoting high-quality development across various industries. In China, for instance, the development of new quality productive forces (新质生产力) is significantly enhanced by educational digitalization, which plays a crucial role in building a strong digital nation and driving technological innovation. Therefore, exploring the digital transformation experiences of different countries or regions is essential for global sharing of insights, contributing to a more prosperous and harmonious world.

As technology continues to innovate, the effectiveness of many emerging media in educational applications still needs to be continuously explored and optimized. In summary, educational technology is an indispensable medium and learning channel for the future. Research on how to effectively apply it while avoiding misuse needs continuous updates and efforts to address research gaps. Besides focusing on positive impacts, it is equally important to explore and address negative effects or outcomes in order to develop better strategies for response and prevention.

Author Contributions: Conceptualization, J.-H.Y. and Y.-W.H.; writing—original draft preparation, J.-H.Y., Y.-W.H. and Y.-F.W.; writing—review and editing, J.-H.Y. and. Y.-F.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Fundamental Research Funds for the Central Universities in China (grant number: 2022NTSS52), Beijing Normal University's First-class Discipline Cultivation Project for Educational Science (grant numbers: YLXKPY-XSDW202408, YLXKPY-XSDW202211, YLXKPY-ZYSB202201), and 2024 Beijing Normal University's Teachers' Teaching Development Fund Project (grant number: 2024125).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

List of Contributions

1. Li, J.; Chang, Y.; Liu, S.; Cai, C.; Zhou, Q.; Cai, X.; Lai, W.; Qi, J.; Ji, Y.; Liu, Y. Higher education in china during the pandemic: Analyzing online self-learning motivation using Bayesian networks. *Sustainability* **2024**, *16*, 7330. https://doi.org/10.3390/su16177330.
2. Zhou, P. Make lectures match how we learn: The nonlinear teaching approach to economics. *Educ. Sci.* **2024**, *14*, 509. https://doi.org/10.3390/educsci14050509.
3. Balanyà Rebollo, J.; De Oliveira, J.M. Identifying strengths and weaknesses in mobile education: A gender-informed self-assessment of teachers' use of mobile devices. *Appl. Syst. Innov.* **2024**, *7*, 31. https://doi.org/10.3390/asi7020031.
4. Wang, K.; Wang, T.; Wang, T.; Cai, Z. Research on evaluation methods for sustainable enrollment plan configurations in Chinese universities based on Bayesian networks. *Sustainability* **2024**, *16*, 2998. https://doi.org/10.3390/su16072998.
5. Zheng, L.; Lee, K.C. Examining the effects of "Small private online course and flipped-classroom"-Based blended teaching strategy on first-year English-major students' achievements. *Sustainability* **2023**, *15*, 15349. https://doi.org/10.3390/su152115349.
6. Bahroun, Z.; Anane, C.; Ahmed, V.; Zacca, A. Transforming education: A comprehensive review of generative artificial intelligence in educational settings through bibliometric and content analysis. *Sustainability* **2023**, *15*, 12983. https://doi.org/10.3390/su151712983.
7. Zhu, F.; Xu, B. The role of dynamic geometry software in teacher–student interactions: Stories from three Chinese mathematics teachers. *Sustainability* **2023**, *15*, 7660. https://doi.org/10.3390/su15097660.
8. Wang, K.; Sang, G.Y.; Huang, L.Z.; Li, S.H.; Guo, J.W. The effectiveness of educational robots in improving learning outcomes: A meta-analysis. *Sustainability* **2023**, *15*, 4637. https://doi.org/10.3390/su15054637.
9. Finger, A.; Groß, J.; Zabel, J. Plant identification in the 21st century—What possibilities do modern identification keys offer for biology lessons? *Educ. Sci.* **2022**, *12*, 849. https://doi.org/10.3390/educsci12120849.

10. Dajani, D.; Yaseen, S.G.; El Qirem, I.; Sa'd, H. P Predictors of intention to use a sustainable cloud-based quality management system among academics in Jordan. *Sustainability* **2022**, *14*, 14253. https://doi.org/10.3390/su142114253.
11. Yamano, H.; Park, J.J.; Choe, N.H.; Sakata, I. Understanding students' perception of sustainability: Educational NLP in the analysis of free answers. *Sustainability* **2022**, *14*, 13970. https://doi.org/10.3390/su142113970.
12. Koutsantonis, D.; Koutsantonis, K.; Bakas, N.P.; Plevris, V.; Langousis, A.; Chatzichristofis, S.A. Bibliometric literature review of adaptive learning systems. *Sustainability* **2022**, *14*, 12684. https://doi.org/10.3390/su141912684.
13. De Alwis, A.C.; Andrlić, B.; Šostar, M. The influence of E-HRM on modernizing the role of HRM context. *Economies* **2022**, *10*, 181. https://doi.org/10.3390/economies10080181.

References

1. Haleem, A.; Javaid, M.; Qadri, M.A.; Suman, R. Understanding the role of digital technologies in education: A review. *Sustain. Pperations Comput.* **2022**, *3*, 275–285. [CrossRef]
2. Burbules, N.C.; Fan, G.; Repp, P. Five trends of education and technology in a sustainable future. *Geogr. Abil.* **2020**, *1*, 93–97. [CrossRef]
3. Crompton, H.; Bernacki, M.; Greene, J.A. Psychological foundations of emerging technologies for teaching and learning in higher education. *Curr. Opin. Psychol.* **2020**, *36*, 101–105. [CrossRef] [PubMed]
4. Sweller, J. Cognitive load theory and educational technology. *Educ. Technol. Res. Dev.* **2020**, *68*, 1–16. [CrossRef]
5. Criollo-C, S.; Guerrero-Arias, A.; Jaramillo-Alcázar, Á.; Luján-Mora, S. Mobile learning technologies for education: Benefits and pending issues. *Appl. Sci.* **2021**, *11*, 4111. [CrossRef]
6. Chan, C.K.Y. A comprehensive AI policy education framework for university teaching and learning. *Int. J. Educ. Technol. High. Educ.* **2023**, *20*, 38. [CrossRef]
7. Ye, J.-H.; Lee, Y.-S.; He, Z. The relationship among expectancy belief, course satisfaction, learning effectiveness, and continuance intention in online courses of vocational-technical teachers college students. *Front. Psychol.* **2022**, *13*, 904319. [CrossRef] [PubMed]
8. UNESCO. Guidance for Generative AI in Education and Research. 2024. Available online: https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research (accessed on 10 August 2024).

sustainability

MDPI

Article

Leveraging Educational Technology in Liberal Arts Dance Sports: Exploring Effectiveness and Sustainable Application

Sunghae Park [1] and Yeonjae Kim [2,*]

[1] Department of Exercise Rehabilitation, Gachon University, Seongnam 13120, Republic of Korea; psh8503@gachon.ac.kr
[2] Interdisciplinary Course for Studies of Arts, Sungkyunkwan University, Seoul 03063, Republic of Korea
* Correspondence: hyun3045@naver.com

Abstract: This study explored the effectiveness and sustainable application of educational technology in a university liberal arts dance sports class. A qualitative case study was conducted with nine students enrolled in liberal arts physical education classes at S University in Seoul in 2023. It focused on how educational technology, including apps such as Pose & Tracking and Slopro, can enhance students' understanding of dance movements, promote active discussion, and facilitate objective assessment. The findings indicate that the use of these technologies not only improved the students' learning experiences by facilitating visualization and feedback but also supported collaborative learning and creative expression. This study highlights the potential of educational technology to enhance the quality of liberal arts physical education by making learning highly engaging and effective.

Keywords: educational technology; dance sports; liberal arts physical education; qualitative case study; effectiveness and sustainable application

Citation: Park, S.; Kim, Y. Leveraging Educational Technology in Liberal Arts Dance Sports: Exploring Effectiveness and Sustainable Application. *Sustainability* **2024**, *16*, 8491. https://doi.org/10.3390/su16198491

Academic Editor: Hao-Chiang Koong Lin

Received: 25 August 2024
Revised: 27 September 2024
Accepted: 27 September 2024
Published: 29 September 2024

1. Introduction

The use of various technologies in education has been transforming learning methods and experiences. Learners can learn anytime and anywhere using various devices such as smartphones or tablets, and they can easily understand abstract and complex concepts through simulation and visualization by using new technologies such as virtual reality (VR) and augmented reality (AR) [1]. These educational technologies not only increase knowledge accessibility and provide easy access to a variety of learning resources but also promote students' creative thinking and problem-solving skills; hence, they play an important role in strengthening the digital technology capabilities required in future society [2].

Educational technology has also led to innovative changes in the field of physical education. The use of educational technology in physical education provides a great opportunity to complement traditional physical education methods and offers learners a more effective learning experience [3]. In the past, the application of education technology in physical education mainly involved video shooting and analysis and access to video learning materials through the Internet. These techniques provided an opportunity for learners to compare and improve their motor skills and maximized the learning effect through real-time feedback [4]. Recently, innovative technologies, such as VR, AR, and wearable technologies, have been applied in physical education classes to further improve the learning experience of students [5].

Research related to educational technology in the field of physical education is being actively conducted mainly in the areas of school and elite sports. In particular, school sports research focuses on educational skills in school sports and physical education classes. Relevant studies mainly emphasize the importance of improving students' motor skills and physical activities and explore methods for improving and evaluating learners' motor skills through video analysis, web-based data access, and learning management systems [6–8]. Recently, the research on the applicability of simulation technology using VR and AR in physical education classes has expanded [9]. Meanwhile, elite sports research relates to elite physical education and player development. In the field of elite physical education, the research on educational technology aims to improve players' physical abilities and technology use [10]. Hence, research efforts are directed toward monitoring and improving the performance of players in real time through wearable devices and data analysis technology using sensor technology [11–13]. In addition, existing studies have explored methodologies for analyzing game videos using video analysis technology and strategically improving player training [14].

Educational technology research in the field of physical education mainly focuses on the development of educational apps, but these technology development studies often have limitations in terms of sustainability, as applications can only be used in certain types of classes [15]. In addition, educational technology research is typically conducted in the areas of school and elite sports, and only a few studies have been carried out in the liberal arts field, which is a part of physical education. Liberal arts and physical education play an important role in educating students about the importance of healthy lifestyles and physical activities; promoting students' physical, emotional, and social development; and enhancing academic achievement through various athletic activities [16]. At the university level, liberal arts and physical education not only promote adult learners' holistic development through physical activities but also create the foundation for their participation in leisure activities and daily sports after graduation, thereby improving their quality of life and leading them toward lifelong sports [17].

Considering these insights, the current study selected a university liberal arts sports program to verify the effectiveness of physical education using educational technology. Specifically, this study explored the effectiveness of using educational technology in university liberal arts classes to address the gap in educational technology research in the field of physical education and the changes in the general education field. To contribute to the understanding of the sustainable application of educational technology in physical education, this study aimed to identify the practical and continuous use of educational technologies in liberal arts physical education classes by presenting class cases involving the use of commercialized app technology. In summary, the purpose of this study is to propose the application of educational technologies in liberal arts dance sports classes while considering effectiveness and sustainable application and to explore ways to increase the efficiency of liberal arts physical education.

2. Literature Review

Educational technology is an academic field that utilizes various technologies to enhance the effectiveness and efficiency of the learning process. Physical educational technology aims to maximize outcomes through various tools and platforms, with the goal of improving learners' motor skill acquisition and physical fitness [4]. Initially, it focused on behaviorist learning theories that aimed to induce behavioral changes in learners. Recently, however, cognitive learning theories have been applied, emphasizing learner-centered approaches that actively engage students in acquiring skills through experience. Immersive technologies such as VR and AR are prime examples of how cognitive learning theories are being applied in physical education [5].

This application is particularly prominent in schools for maximizing students' motor skills and participation. Examples of important tools are video analysis, web-based resource access, and learning management systems [18]. Students can visually review their performance and receive feedback, providing a shift from a teacher-centered process to a learner-centered one [19]. Recently, there has been a rise in the use of VR and AR, as they simulate real exercise environments [20]. These technologies provide students with diverse exercise experiences in a safe environment, combining theoretical learning with practical application.

Beyond school physical education, advancements in educational technology have also brought significant changes to the training methods in elite sports and are used to enhance athletes' performance. Sensor technology and wearable devices enable real-time monitoring of athletes' physical conditions, with the collected data being analyzed to provide personalized feedback and enhance performance [21,22]. Additionally, video analysis technology is used to meticulously analyze movements during competitions, and based on this, training programs are strategically adjusted [23]. These technologies not only contribute to improving athletes' skills but also play an important role in injury prevention [14].

Despite the great potential for educational technology in liberal arts physical education, research on its use in this field remains relatively scarce. This application plays an important role in promoting students' holistic development and fostering healthy lifestyle habits [24]. It is particularly crucial for adult learners, providing them with a foundation to maintain health by continuing to engage in physical exercise [25]. This includes personalized feedback through technology, analysis of exercise data, and continuous activity tracking via online platforms [26]. These activities have significant potential for more effective application in general physical education. Therefore, research in this area is needed to extend knowledge of the potential benefits of the use of educational technology and to enhance student engagement and learning outcomes.

Research investigating the potential use of educational technology in physical education can expand toward utilizing commercially available technologies for sustainable application. Existing research has mainly focused on the development of educational apps, often restricted to specific class formats or environments, limiting the broader application of these technologies in physical education [18]. However, by utilizing easily accessible commercially available technologies, these limitations can be overcome. Practical methods can be explored that can be applied sustainably. In particular, research that proposes sustainable educational methods based on commercially available apps or technologies is valuable, as it distinguishes itself from the focus on technology development research.

3. Research Method

3.1. Research Design

A qualitative case study method was used as the research method for investigating the effectiveness and sustainable application of educational technology in a liberal arts dance sports physical education class. Stake (2005) [27] and Denzin and Lincoln (2005) [28] describe the case study as a qualitative research method that explores a specific case or phenomenon in depth. It typically focuses on individuals, groups, organizations, or events, with the researcher analyzing the case from multiple perspectives to understand its context, process, and outcomes. Case studies are typically characterized by a detailed examination of a small number of cases, with the primary focus on explaining and understanding phenomena rather than verifying specific theories. The current study found the qualitative case study method as suitable for finding meaning within the overall context of education [29] to assess the effectiveness and sustainability of educational technology in liberal arts physical education classes.

3.2. Researchers and Participants

Researcher A is a Ph.D. in Sports Pedagogy with experience in physical education instruction at various universities since 2014. She directly taught the liberal arts dance sports course, which is the subject of this study, applying educational technology. Researcher B is also a Ph.D. in Sports Pedagogy and has taught physical education and dance education at several universities since 2015. She observed the liberal arts dance sports course incorporating educational technology and provided feedback on the class. The educational philosophy, objectives, and content organization for the course design in this study were developed based on the close collaboration between the two researchers.

The participants in this study were students enrolled in the liberal arts dance sports course at S University in Seoul, which was offered during the first semester (2 classes) and the second semester (1 class) of 2023. S University offers a wide range of elective physical education classes each semester and is well-equipped with facilities suitable for various physical activities. Among these, the liberal arts dance sports course is popular and attracts students from various majors. Each class consists of 30 students, with 60 students enrolled in the first semester and 30 students in the second semester, totaling 90 students. To ensure active participation and engagement, fostering students' interest in dance sports is essential.

The main research participants were nine students who actively and consistently participated in the classes without absences and agreed to take part in the interviews. These participants were selected using a purposive sampling method to ensure that they met specific criteria relevant to the study [30]. Although the sample size is small, the data collected from these nine participants were sufficient to achieve the research objectives. Reflection papers and interviews were conducted at three points during the semester—at the beginning, middle, and end—allowing for a comprehensive collection of data. The repeated collection of data through both reflection papers and interviews provided in-depth insights into the effectiveness and sustainability of the educational technology applied in the course. Therefore, despite the small sample size, the selected participants and the robust data collection process ensured that the study's goals were adequately met. Detailed participant information is shown in Table 1.

Table 1. Participant information.

Year	Semester	Participants	Age	Gender	Major	Previous Elective Physical Education Experience
2023	1st Semester (2 classes)	1	21	Male	Business	Aerobics
		2	20	Female	Psychology	-
		3	22	Male	Computer Science	Badminton, tennis
		4	25	Male	Design	-
		5	23	Male	Economics	-
	2nd Semester (1 Class)	6	20	Female	Economics	-
		7	22	Male	Physics	Basketball, soccer
		8	22	Female	Sociology	Aerobics
		9	21	Female	Chemistry	-

3.3. Data Collection

Data were collected from a literature analysis, interviews with student participants, students' reflection papers, and the instructor's reflection journal. First, during the lesson planning phase, a theoretical review of educational technology was conducted to carefully examine which educational technology was suitable for the dance sports classes. During the course, data, including photographs and video recordings related to the use of educational technology, were collected.

Second, the researchers analyzed the students' reflection papers and conducted interviews. The collection of reflection papers and the interviews were conducted three times: at the beginning, middle, and end of the semester. The interview questions were as follows. At the beginning of the semester, the questions were about the students' reasons for enrolling in the class and their expectations of the course. In the middle of the semester, the questions were about interim evaluations of the course and assessments of the educational technology used up to the eighth session. At the end of the semester, the questions were about the overall evaluation of the course, the effectiveness of the educational technology used, and whether it would be beneficial to continue using the technology in the future.

Finally, the instructor's reflection journal was collected. Over the course of 15 sessions, the journal documented observations of students' changes, the instructor's own process of change, and the strengths and weaknesses of the operational aspects of the class. This journal was utilized as part of the research data. Additionally, the two researchers distinguished between the roles of instructor and observer and conducted discussions before and after each session to enhance teaching effectiveness.

3.4. Data Analysis

This study aimed to explore the effectiveness of educational technology by uncovering the meaning of learners' experiences conveyed through the collected data as qualitative research results [31]. In particular, it sought to understand the meaning revealed by the various materials collected during class with respect to the application of educational technology. Text materials related to class experiences (reflection papers, interviews, and the instructor's reflection journal) were analyzed through description, analysis, and interpretation [32].

The data analysis was conducted in three main stages. First, various data were collected and systematically organized, followed by a coding process to identify key themes. In the analysis stage, a comparative analysis of the data was conducted based on the identified themes, interpreting their meaning within a broader context. In the final interpretation stage, the data were analyzed considering the specific context of each case, which led to answering the research questions. The research findings were described based on the key findings, revealing the positive impact of educational technology on the learning process.

To enhance the validity of the data analysis, member checks were conducted, and peer debriefing was carried out with teacher experts. Peer debriefing involved consultations with two experts holding doctorate degrees in physical education and dance education, one university physical education lecturer, and two dance instructors to minimize subjective errors [33].

4. Course Design and Management

The course, which incorporated educational technology, consisted of three liberal arts dance sports classes conducted from March to December 2023. The liberal arts dance sports course at S University is one of the most popular among various liberal arts physical education courses, with registration closing early due to high demand and a class size limit of 30 students. These liberal arts physical education courses mainly target college students who are adept at handling various digital technologies; hence, they are an appropriate venue in which to apply educational technology. Unlike elite sports, liberal arts physical education courses cater to students with diverse majors. Therefore, the primary objective is not to teach sports skills but to emphasize the importance of healthy lifestyle habits and physical activity through various exercise programs using educational technology. The structure and operation of the dance sports course are listed in Table 2.

In Lessons 2 to 7, the students focused on learning the basic postures and movements to systematically learn the functions of dance sports. The educational technologies used in this process were Pose & Tracking and Slopro (see Figure 1). The Pose & Tracking app visualizes postures and movement trajectories recorded on video. Hence, by using the Pose & Tracking app, students can correct their postures and movements to meet the requirements of dance sports. This app allows students to learn dance sports poses anytime and anywhere and helps them practice even without a partner. Slopro is an educational technology useful for learning an entire movement. Slopro is a tool for editing videos at various playback speeds and applying any speed to different parts of a video. Hence, students can use Slopro to play dance sports videos in slow motion, enabling them to focus and learn challenging movements more effectively.

Table 2. Content and activities of liberal arts dance sports class using educational technology.

Lesson	Class Content	Key Activities	Educational Technologies
1	Orientation	Introduction of instructor and overview of the course	
2	Learning basic movements	Practice basic dance sports posture and movement	**Pose & Tracking: Visualizes** posture and movement trajectories Slopro: Edits video playback speeds for slow-motion viewing
3			
4			
5			
6			
7			
8	Midterm	Practical test on basic dance sport movement	
9	Creating choreographic compositions	Partner choreography	**Online System**: Provides discussion spaces for idea development **Living Archive**: AI-generated movements for inspiration
10			
11		Group choreography	
12			
13	Presenting team projects	Midterm review of group creations	**VLLO**: Video editing app
14		Group creation presentations	
15		Class warm-up and discussion of the presentations	

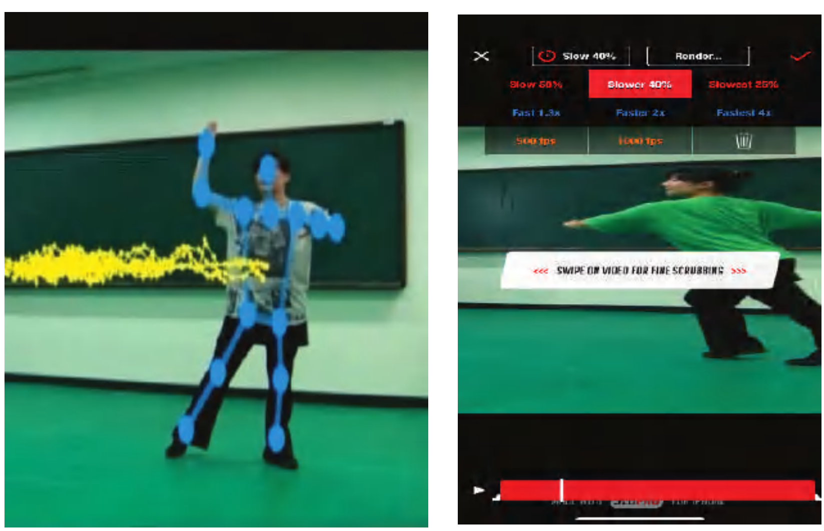


Figure 1. Examples of Pose & Tracking and Slopro used in liberal arts dance sports.

Lessons 9–12 involved creating various choreographic compositions using the basic movements of dance sports. This creative process helps learners develop a deeper understanding of dance sports movements and fosters creative thinking. The first part of the process involved choreography involving student pairs. Subsequently, the process expanded to group choreography to reduce the students' apprehension about creating. During the choreography classes, the university's online system was actively used to provide discussion spaces, through which the students could refine their ideas by engaging in various discussions related to choreography both during and after class. Additionally, the Living Archive program was proposed. The Living Archive program uses AI deep learning technology to generate movements. By utilizing the movements created in the Living Archive, students can gain ideas about their choreography and easily construct their compositions.

Finally, Lessons 13–15 involved presenting the works created by each team. Expanded from the pairs in the first part of the process, each team comprised four to six members. Team formation was voluntary. Instead of presenting their work offline, the teams submitted their presentations as videos. These videos were compiled using video editing apps or programs such as VLLO. The team projects were to be a single cohesive piece with a story centered on one theme and feature freely composed and edited content. This approach allowed the students to experience the creative process and express their ideas innovatively using various editing techniques. This method is particularly suitable for the current generation, who are accustomed to smartphones and expressing their thoughts using videos. Additionally, the learners experienced a unique aspect of the class by engaging in the creative process; this feature differentiated the liberal arts dance sports course from others in which dance is taught alone. The first presentation of the work can be shown offline or as a flow using a presentation. In the final lesson (Lesson 15), the videos submitted by each team were viewed together, followed by a critique and discussion session.

5. Results

5.1. Effectiveness of Educational Technology in Understanding Movements

The application of educational technology in liberal arts dance sports classes contributes to enhancing students' understanding of dance sports content, particularly dance movements. The students who participated in the dance sports classes applying educational technology explained that the technology used in the class helped them learn basic dance sports movements.

> *I particularly liked the fact that I used the app to learn dance sports postures and movements at the beginning of the semester. When the professor explained the movement, she gave me various feedback such as "This is a good posture to dance in. Adjust the size of your stride in this way," but it is difficult to check how I am dancing because it disappears once you dance. However, after visualizing my posture and movement using the app, I understood the feedback of my professor better(Student 3, 2023 Spring Semester, Class A, Interview).*

Student 3 mentioned that they had positive experiences in the process of learning posture and movements using an app in the early stages of learning dance sports. As dance movements disappear over time, students cannot immediately check their movements [34]. In this study, the dance sports classes using apps helped the students understand the instructor's feedback more clearly by visualizing their movements in real time. In other words, students can objectively observe their appearance and improve their posture and movement using an app that visualizes their posture and movement.

The students also said that the educational technology adopted in the dance sports classes was useful in learning difficult movements.

> *I used the Slopro app really often. It was a great help to me because I was able to learn intensively while slowly and repetitively playing difficult movements. It also relieved the pressure to memorize all the steps in class, so I was able to more enjoyably participate in the class (Student 4, 2023 Spring Semester, Class B, Class Reflection Paper).*

Student 4 mentioned that he frequently used the Slopro app in his class reflection paper to support his learning. One of the core functions of the Slopro app is its ability to play movements slowly, thus enabling learners to slowly analyze complex or difficult movements and learn them repeatedly. This approach is based on the principle of chunked learning, in which information is divided into small units [35]. In other words, through the slow reproduction function of the app, the students could analyze complex dance sports movements in small parts and learn them repeatedly. This process helped them learn the detailed elements of the entire movement step by step. One of the greatest advantages of chunked learning is that it reduces the cognitive burden on learners, as mentioned by Student 4 at the end of the reflection journal. Dividing the information into small units allows each unit to be learned individually, allowing learners to understand and remember complex information more effectively and reducing the cognitive burden [36].

Educational technology not only helps learners understand their own movements but also plays a significant role in understanding their partner's movements.

> *In my case, if I used the app to learn my movements at the beginning of the semester, I think I would do so to understand my partner's movements, especially when I was working with him on creative activities. In dance sports, my partner and I must dance as one body, so I need to know how my partner is moving so I can dance well accordingly (Student 7, 2023 Fall Semester, Interview).*

Student 7 explained that at the beginning of the semester, the app was used to focus on individual movement practice and modification; however, in the second half, it was used to understand and coordinate the partner's movements. Dance sports are cooperative partner activities, and the understanding and interdependence of partners are important factors [37]. By the middle and second half of the semester, the students recognized the need to coordinate and match their movements through collaboration with their partners. At this time, the focus was on visually analyzing and coordinating their partners' movements using the app. In other words, because the coordination of movements with partners is important in dance sports [38], the students performed active learning to analyze their partners' movements in real time and adjust them through the app. Accordingly, the app functioned as a tool for cooperative learning and interactive feedback during the middle and second half of the semester [39].

In summary, the application of educational technology in liberal arts dance sports classes helps students effectively understand feedback through the visualization of movements. It also reduces the learning burden through segmental and repetitive learning and facilitates effective collaboration with partners.

In addition to these key benefits, educational technology played an important role in enhancing the students' creative thinking and autonomy in learning. By utilizing technology, the students were able to experiment with and modify various movements, thereby fostering their creativity. They also gained control over their own learning processes and made improvements accordingly. Furthermore, by comparing their initial movements with revised ones through real-time feedback, the students were able to strengthen their motivation for learning. As a result, educational technology went beyond merely being a tool for learning movements—it promoted creativity, self-directed learning, and motivation, thereby improving the students' overall learning experience.

5.2. Effectiveness of Educational Technology in Discussion and Assessment

In this study, the application of educational technology in the dance sports classes provided various advantages in terms of understanding the content of dance sports and facilitating discussion and evaluation. The students explained that they were able to engage in more active discussions through the educational technology used in the class.

> *I have to use various apps in class, so I have to record the movements every time. I was able to participate in the discussion more actively based on the recorded movements in the creative class with a partner and team starting from the middle of the semester. In particular, it was great to record past movements in the creative class at the end of the semester. If I don't remember what movement I did last time, I can just watch a recording of it (Student 9, 2023 Fall Semester, Interview).*

Student 9 mentioned that creative dance sports activities using movement records led to active participation in the discussions. As recorded data increased the concreteness and reliability of discussions, the students could clearly communicate their ideas and feedback by recording movements and conducting discussions with the team [40]. By referring to the recording of movements, team members shared more specific and practical feedback, which helped solve problems during the creative process [41]. Furthermore, the recording of movements provided the students with the opportunity to review their previous movements. By keeping records of past movements, the students could continuously refer to the learning content, which aided their memory.

From the instructor's perspective, the fact that students who were absent from class could also participate in the discussion through the use of educational technology served as a great advantage.

> *I was wondering how students who are absent or sick should participate in the creative process every semester, but because the movements created by the students can be shared on both sides through eTL (university online learning system) while recording them, the students who are absent can also be included in the discussion, thereby increasing their overall participation in the creative process(Researcher A, Reflection Journal).*

Students who are absent must have access to information and materials to participate in creative activities. By recording and sharing the recorded movements through eTL, students who are absent can visually check their progress and understand the context of their creative activities. This feature provides an opportunity for students to be included in the discussion and creative process even when they are not in class. It also allows them to continue their learning experience by minimizing the learning gap due to absence.

The educational technology used in the dance sports classes not only enabled active discussion but also facilitated objective evaluations for the students.

> *It was good to be able to double-check whether I had properly evaluated the recorded movements during the peer evaluation. Evaluation should be done objectively, but movements have one-off characteristics, so if there were no recorded movements, it would have been difficult to confirm whether what I evaluated was real or wrong (Student 1, 2023 Spring Semester, Class A, Class Reflection Paper).*

Student 1 mentioned that the objectivity and reliability of peer evaluation could be increased by reviewing recorded movements through the educational technology. As a movement is one-off and instantaneous because of its nature, evaluators have no choice but to rely on simple memory to evaluate movements [42]. In the dance sports class, the recorded creations were submitted using the adopted educational technology. Hence, the evaluator could review the movements repeatedly as necessary. This feature was an opportunity to minimize errors in the evaluation process and increase its accuracy.

In summary, the application of educational technology in liberal arts dance sports classes has advantages, as it encourages active discussion among students, minimizes learning gaps for absent students, and enables objective evaluation.

These technologies provided the students with the opportunity to reflectively analyze both their own and others' movements, contributing to the enhancement of self-evaluation skills. Additionally, by reviewing recorded movements in real-time, the students developed critical thinking skills and created an inclusive learning environment where absent students could participate in class. During peer evaluation, the ability to repeatedly review recorded movements also improved the fairness and accuracy of assessments. Consequently, educational technology played a crucial role in strengthening the students' critical thinking, inclusivity, and fairness in the evaluation process.

5.3. Effectiveness of Educational Technology in Expression and Creation

In liberal arts dance sports classes, educational technology offers various advantages in terms of student expression and creation. In this study, the students explained that the educational technology used in class helped them express and create movements.

> *It was very nice to be able to express what I learned in dance sports in various ways through the creative process. In the previous class, I only understood the knowledge I had learned, but in this class, I use and apply what I have learned through creation to express*

it in various ways. In particular, what I liked about this class was that I could try various movements using various techniques to apply the knowledge I learned(Student 5, 2023 Spring Semester, Class B, Interview).

Student 5 favored the creative process in class that allowed them to apply their knowledge. The student also highlighted the possibility of trying various movements using various techniques. The effectiveness of learning is revealed when students are able to fully utilize and apply the knowledge they have acquired in learning [43]. In this respect, the creative process in the dance sports classes was important because it encouraged the students to use what they had learned in a practical and creative manner [44]. An important note in the interview content of Student 5 is that the various educational technologies used in the class provided an opportunity to experiment with and explore various expression methods in the creative process. In other words, educational technology played a role in opening the possibility for students to explore new expression methods beyond traditional techniques.

Meanwhile, educational technology also played a role in reducing the burden of creation and enhancing the positive experience of creation for students who faced difficulties in the process.

As a non-major, it was difficult for me to express or create movements, but it was easier to use the Living Archive program introduced by the professor. Rather, the process of creating felt more interesting and enjoyable(Student 8, 2023 Fall Semester, Class Reflection Paper).

Student 8 explained in the class reflection paper that the Living Archive program helped alleviate difficulties faced in the creative process. Non-majors are likely to feel inherently burdened and experience difficulties expressing or creating movements [45]. This burden is related to a lack of related skills or knowledge, unfamiliarity with the creative process, or lack of confidence resulting from comparisons with experts [46]. The fact that the students felt that the creative process was interesting and enjoyable through the Living Archive program means that the educational technology provided an environment where the students, the subjects of the creative process, could feel satisfied and experienced motivation. In other words, the support and resources provided by the Living Archive program in the creative process positively changed the creative experiences of non-majors.

In summary, the application of educational technology in liberal arts dance sports classes encourages students to express various movements and positively changes the creative experiences of non-majors.

Educational technology served as an important tool in fostering the students' creative expression. Through various technologies, the students were able to explore movements in new ways beyond traditional methods, allowing them to apply the knowledge they had learned in more practical and creative ways. Additionally, it supported self-directed learning, enabling the students to progress at their own pace during the creative process and explore more options. For non-major students, the technical support helped reduce the burden they felt during the creative process and increased their motivation. This enhanced their confidence and satisfaction, allowing them to become more immersed in the creative process and have a more positive learning experience. Consequently, educational technology supported the students' learning experiences in multiple aspects, including creativity, self-directed learning, and motivation, significantly impacting their learning process.

6. Discussion

This study utilized various educational technologies in liberal arts dance sports classes to explore their effects. Based on the main results of this study, the discussions related to the application of educational technology in liberal arts physical education are as follows.

First, the application of educational technology in liberal arts physical education is vital in the visualization of movements. Traditional physical education methods tend to require instructors to demonstrate or rely on oral explanations for the accurate execution of movements [47], which change and dissipate over time [24]. Therefore, the visualization of the movement process is essential for effective understanding and learning. Prior studies related to educational technology in the fields of school sports or elite sports have also emphasized the visualization of movement using educational technology [10]. Similarly, this study examined the application effects of educational technology in liberal arts sports and found that the use of the technology significantly improved the learners' understanding and learning experiences. Visualization of movement through educational technology contributed to enhancing the understanding of the behavior of individual and paired students, provided an environment in which learning evaluation and discussion could be conducted more effectively, and supported the expression and creation processes. These results suggest that the use of technology in physical education maximizes the educational effect by facilitating the visualization of movement and movement processes. This characteristic of immediate feedback in physical education through technology can contribute to improving the quality of physical education.

Second, the application of educational technology plays an important role in improving students' self-awareness, self-regulation, and self-evaluation as metacognitive strategies in liberal arts physical education. Metacognitive strategies enable learners to self-check and control their cognitive processes and learning methods, including learning processes, understanding, and problem solving [48]. In physical movement classes, self-awareness, self-regulation, and self-assessment are key factors for improving movement skills and learning efficiency, which have important implications for feedback and self-reflection related to movement performance [49]. This study presents a case in which various educational technologies used in liberal arts dance sports classes effectively support this metacognitive strategy. Using a video app, the students were able to recognize their movements and modify, analyze, and evaluate them according to their ideal posture. Additionally, these technical tools allowed the students to learn freely, actively participate anytime and anywhere, and experience the recognition, regulation, and evaluation processes. These technical tools contributed to the students' clearer understanding of their movements, systematically improved movement, and enabled objective evaluation. Therefore, the use of educational technology in liberal arts physical education is useful in strengthening learners' metacognitive skills and enriching learning experiences.

However, despite the positive effects explored in this study, there are several limitations to applying educational technology in liberal arts physical education. The disparities in technological proficiency, accessibility issues, and complexity of tool usage can hinder the sustainable use of educational technology. For example, differences in students' abilities to use technology may lead to learning gaps, and unequal access to devices or the Internet could result in some students being excluded from learning opportunities. Additionally, overly complex tools can reduce usability, ultimately lowering learning efficiency. To address these challenges and ensure the sustainable use of educational technology, it is essential to provide technical training and support programs for both instructors and students. Furthermore, improving access to devices and resources, as well as selecting intuitive and efficient tools, is crucial. Establishing a system for continuous monitoring and feedback on the use of educational technology will further enhance its long-term effectiveness and sustainability.

Third, this study is noteworthy as it explored the sustainable application of commercialized and easily accessible educational technologies in liberal arts physical education. As mentioned in the Introduction, the research on educational technology in the field of physical education mainly focuses on the development of educational apps, and technology development studies are limited to specific types of instruction [15]. The current study overcomes these limitations and presents practical educational methods using existing commercial app technologies. As a result of using various commercialized technologies in liberal arts dance sports classes, the changes in the way the students effectively used the technologies could be observed. Educational technology enriched learners' learning experiences and supported each stage of learning more effectively through technical tools ranging from visualizing movements and understanding feedback in the early stages, coordinating and facilitating cooperative learning in the middle stage, and promoting evaluation and creative processes in the second half. The accessibility and practicality of the technologies played an important role in making the students participate more actively in learning and effectively integrate technology, thereby contributing to increasing the efficiency of education [50]. As a generation familiar with technology, the current generation of learners has high digital utilization capabilities and is highly interested in teaching methods that use technology [51]. Therefore, the sustainable use of educational technology must be considered to induce in-depth participation in liberal arts physical education.

7. Conclusions and Recommendations

This study aimed to analyze the effectiveness of applying educational technology in liberal arts dance sports classes and explore sustainable application strategies through case studies. The use of educational technology in the liberal arts dance sports classes significantly enhanced the students' educational experiences. With the incorporation of educational technology, the students received clearer and more immediate feedback, which improved their understanding of movement techniques. In addition, the application of educational technology increased student engagement through active discussions, reduced learning gaps among absent students, and provided objective assessment methods. Furthermore, the use of educational technology exerted a positive impact on the students' creative expression and creative processes.

To further enhance the effectiveness and sustainable application of educational technology in liberal arts physical education, the following recommendations are proposed. First, faculty members must receive specialized training in the functions and utilization of educational technology tools available for physical education so that they can effectively integrate these technologies into their teaching. Second, the accessibility of feedback systems and learning materials supported by educational technology should be improved to ensure that students fully benefit from these tools. Third, efforts must be made to establish a system for continuously monitoring and evaluating student engagement and learning outcomes when integrating educational technology into physical education classes to promptly address and incorporate the necessary improvements. Fourth, platforms for sharing and discussing case studies related to educational technology applications in various physical activities should be created to support faculty members in exploring more creative and effective ways to use technology. These approaches will contribute to enhancing the quality of liberal arts physical education and improving students' learning experiences.

In addition to these practical recommendations, future research should consider the rapidly evolving nature of educational technology and its potential applications in diverse educational settings. As educational technologies such as virtual reality (VR), augmented reality (AR), and artificial intelligence (AI) continue to advance, their use in physical education is expected to expand beyond traditional methods. Future studies should explore how these emerging technologies can be integrated into liberal arts physical education to further enhance students' engagement, learning outcomes, and creativity. Moreover, research should address the potential challenges posed by technological advancements, such as disparities in access, technological proficiency, and the sustainability of integrating rapidly changing tools into the curriculum. Investigating how to balance cutting-edge technology with accessibility and usability will be crucial in ensuring equitable learning opportunities for all students. Lastly, collaboration between educators, technology developers, and policymakers will be essential to establish frameworks that can support the long-term and adaptable use of educational technology in physical education, accommodating its dynamic nature.

These future-oriented research directions will not only ensure the continued effectiveness of educational technology in liberal arts physical education but also help educators stay ahead of technological trends and adapt to the changing landscape of educational environments.

Author Contributions: Conceptualization, S.P. and Y.K.; methodology, S.P.; validation, Y.K.; data curation, S.P. and Y.K.; writing—original draft preparation, S.P.; writing—review and editing, Y.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived for this study because student feedback used in this paper was part of regular and teaching and learning activities.

Informed Consent Statement: Informed consent was obtained from all subjects involved in this study.

Data Availability Statement: The data are not publicly available due to privacy issues.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of the data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Hillery, A. Data Literacy Is the Key to the Future of Data Visualization. Nightingale, October 2019. Available online: https://medium.com/nightingale/why-data-literacy-is-the-key-to-the-future-of-data-visualization-57639532377f (accessed on 24 August 2024).
2. An, T.; Oliver, M. What in the world is educational technology? Rethinking the feld from the perspective of the philosophy of technology. *Learn. Media Technol.* **2021**, *46*, 6–19. [CrossRef]
3. Julia, S.; Calderón, A. Technology-Enhanced Learning Physical Education? A Critical Review of the Literature. *J. Teach. Phys. Educ.* **2021**, *41*, 689–709. [CrossRef]
4. Ashley, C.; Jones, B. Using Digital Technology to Enhance Student Engagement in Physical Education. *Asia-Pacific J. Health Sport Phys. Educ.* **2011**, *2*, 51–66. [CrossRef]
5. Sunyoung, K.; Kang, S. The Study on the Application of Virtual Reality in Adapted Physical Education. *Clust. Comput.* **2019**, *22*, 2351–2355. [CrossRef]
6. Jorge, L.; Valverde-Esteve, T.; García-Massó, X. Use of Mobile Devices as a Facilitator of the Practice of Physical Activity in Physical Education Lessons: Experience in Higher Education. *J. Phys. Educ. Sport* **2020**, *20*, 3629–3634. [CrossRef]
7. Deborah, L. Next Generation PE'? A Socio material Approach to Digitized Health and Physical Education. *Sport Educ. Soc.* **2021**, *27*, 516–528. [CrossRef]
8. Amanda, M.; Gerdin, G. Challenging Gendered Inequalities in Boys' Physical Education Through Video-Stimulated Reflections. *Sport Educ. Soc.* **2018**, *23*, 761–772. [CrossRef]
9. Michael, M.; Müller, J.; Reimers, A.K. Use of Digital Media for Home-Based Sports Activities During the COVID-19 Pandemic: Results from the German SPOVID Survey. *Int. J. Environ. Res. Public Health* **2021**, *18*, 4409. [CrossRef]
10. Markus, R.; Robert, S.; Charles, A.M. Technology meets sport psychology: How technology and artificial intelligence can shape the future of elite sport performance. *J. Sport Psychol. Action* **2024**, *15*, 63–69. [CrossRef]
11. Muniz-Pardos, B.; Sutehall, S.; Gellaerts, J.; Falbriard, M.; Mariani, B.; Bosch, A.; Asrat, M.; Schaible, J.; Pitsiladis, Y.P. Integration of wearable sensors into the evaluation of running economy and foot mechanics in elite runners. *Curr. Sports Med. Rep.* **2018**, *17*, 480–488. [CrossRef]
12. Fuss, F.K.; Düking, P.; Weizman, Y. Discovery of a sweet spot on the foot with a smart wearable soccer boot sensor that maximizes the chances of scoring a curved kick in soccer. *Front. Physiol.* **2018**, *9*, 63. [CrossRef] [PubMed]
13. Lo Presti, D.; Massaroni, C.; Saccomandi, P.; Caponero, M.A.; Formica, D.; Schena, E. A wearable textile for respiratory monitoring: Feasibility assessment and analysis of sensors position on system response. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2017**, *2017*, 4423–4426. [CrossRef] [PubMed]
14. Ciara, E.; Ross, W.; Karen, H.; Melissa, D. Construction and communication of evidence-based video narratives in elite sport: Knowledge translation of sports injury experiences. *J. Appl. Sport Psychol.* **2023**, *35*, 731–754. [CrossRef]
15. Cao, F.; Lei, M.; Lin, S.; Xiang, M. Application of artificial intelligence-based Big Data AI technology in physical education reform. *Mob. Inf. Syst.* **2022**, *2022*, 4017151. [CrossRef]
16. Ben, D.; Seunghyun, B.; Donal, H.; Yongjin, L.; Judy, F. Students' Perspectives of Social and Emotional Learning in a High School Physical Education Program. *J. Teach. Phys. Educ.* **2024**, *43*, 549–556. [CrossRef]
17. Emma, R.; Lewis, S.; Miah, A.; Lupton, D.; Piwek, L. *Digital Health Generation? YoungPeople's Use of 'Healthy Lifestyle' Technologies*; University of Bath: Bath, UK, 2020.
18. Gil-Espinosa, F.J.; Nielsen-Rodríguez, A.; Romance, R.; Burgueño, R. Smartphone applications for physical activity promotion from physical education. *Educ. Inf. Technol.* **2022**, *27*, 11759–11779. [CrossRef]
19. Cho, J.H.; Song, G.J.; Oh, J.H.; Chung, H.W. Current Mobile App Contents and Technologies for Physical Activity Promoting Intervention. *Korean J. Meas. Eval. Phys. Educ. Sport Sci.* **2016**, *18*, 1–16.
20. Yu, C.W. A Case Study on Physical Education Class Using Physical Activity Solution App. *J. Res. Curric. Instr.* **2022**, *26*, 458–471. [CrossRef]

21. Yang, Z.; Wang, Q.; Yu, H.; Xu, Q.; Li, Y.; Cao, M.; Dhakal, R.; Li, Y.; Yao, Z. Self-Powered Biomimetic Pressure Sensor Based on Mn–Ag Electrochemical Reaction for Monitoring Rehabilitation Training of Athletes. *Adv. Sci.* **2024**, *11*, 2401515. [CrossRef]
22. Guo, E.; Cui, X. Simulation of optical sensor network based on edge computing in athlete physical fitness monitoring system. *Opt. Quantum Electron.* **2024**, *56*, 637. [CrossRef]
23. Kebao, Z.; Kehu, Z.; Liu, W. The Evaluation of Sports Performance in Tennis Based on Flexible Piezoresistive Pressure Sensing Technology. *IEEE Sens. J.* **2024**, *24*, 28111–28118.
24. Mokmin, N.A.M. The effectiveness of a personalized virtual fitness trainer in teaching physical education by applying the artificial intelligent algorithm. *Int. J. Hum. Mov. Sports Sci.* **2020**, *8*, 258–264. [CrossRef]
25. Mokmin, N.A.M.; Jamiat, N. The effectiveness of a virtual fitness trainer app in motivating and engaging students for fitness activity by applying motor learning theory. *Educ. Inf. Technol.* **2021**, *26*, 1847–1864. [CrossRef]
26. Amanda, W.; Jonathan, R.; Scott, H.; Ryan, E.R.; Sam, L. Advancing physical activity research methods using real-time and adaptive technology: A scoping review of 'no-code' mobile health app research tools. *Sport Exerc. Perform. Psychol.* **2024**, 1–19. [CrossRef]
27. Stake, R.E. *Multiple Case Study Analysis*; Guilford Press: New York, NY, USA, 2006.
28. Denzin, N.K.; Lincoln, Y.S. *The Sage Handbook of Qualitative Research*, 3rd ed.; Sage: Thousand Oaks, CA, USA, 2005.
29. Silverman, D. *Qualitative Research*; Sage: Thousand Oaks, CA, USA, 2021.
30. Maggi, S.-B.; Claire, H.M. *Qualitative Research: The Essential Guide to Theory and Practice*; Routledge: Oxfordshire, UK, 2022.
31. Mayring, P. Qualitative Content Analysis. In *Qualitative Research: Handbook*; Flick, E.U., Kardorff, V., Steinke, I., Eds.; Sage: London, UK, 2022.
32. Matthew, B.M.; Michael, H.A.; Johnny, S. *Qualitative Data Analysis: A Methods Sourcebook*; Sage: London, UK, 2019.
33. Creswell, J.W. *Qualitative Inquiry and Research Design: Choosing among Five Approaches*, 2nd ed.; Sage: London, UK, 2007.
34. Eliza, J.G.; Rachael, F.; David, A.W. A Comparison of TAG teach and Video Self-Evaluation to Improve Dance Movement Accuracy. *J. Appl. Behav. Anal.* **2023**, *56*, 914–922. [CrossRef]
35. Brendan, H.; Damien, C. An examination of student preference for traditional didactic or chunking teaching strategies in an online learning environment. *Res. Learn. Technol.* **2021**, *29*, 1–13. [CrossRef]
36. Lah, N.C.; Saat, R.M.; Hassan, R. Cognitive strategy in learning chemistry: How chunking and learning get together. *Malays. Online J. Educ. Sci.* **2014**, *2*, 9–15.
37. Korpinen, K.; Anttila, E. Strange encounters in times of distancing: Sustaining dialogue through integrating language and dance in primary education. *J. Res. Arts Sports Educ.* **2022**, *6*, 45–62. [CrossRef]
38. Barbora, P.U.A.; Romualdas, K.M.; Sarunas, A.S. Effects of Education Programs on Dance Sport Performance in Youth Dancers. *Eur. J. Contemp. Educ.* **2019**, *8*, 136–143. [CrossRef]
39. Hu, Z. Research on the Application of Virtual Reality Technology in the Teaching of Sports Dance in Colleges and Universities. In Proceedings of the 2021 2nd International Conference on Artificial Intelligence and Education (ICAIE), Dali, China, 18–20 June 2021; pp. 414–418. [CrossRef]
40. Maria, P.L.; Arturo, D.S.; Francisco-Jose, B.-B. Students' voices on the effects of expressive activities on their creativity, affectivity, and social relationships. *Res. Danc. Educ.* **2024**, 15. [CrossRef]
41. Canales-Lacruz, I.; Rovira, G. Social Interaction and Group Work in Corporal Expression Activities. *Res. Danc. Educ.* **2020**, *22*, 3–16. [CrossRef]
42. Lu-Ho, H.; Gwo-Jen, H. Enhancing students' choreography and reflection in university dance courses: A mobile technology-assisted peer assessment approach. *Br. J. Educ. Technol.* **2021**, *52*, 266–287. [CrossRef]
43. Huang, C.-E. Discovering the creative processes of students: Multi-way interactions among knowledge acquisition, sharing and learning environment. *J. Hosp. Leis. Sport Tour. Educ.* **2020**, *26*, 100237. [CrossRef]
44. Ørbæk, T. Bodily Learning Through Creating Dance: Student Teachers' Experiences From Norwegian Physical Education Teacher Education. *Front. Sports Act. Living* **2021**, *3*, 758944. [CrossRef]
45. Engdahl, C.; Ceder, S. Exploring Movement in Creative Dance: Introducing 'Dancemblage' in Physical Education Teacher Education. *J. Res. Arts Sports Educ.* **2023**, *7*, 43–58. [CrossRef]
46. Engdahl, C.; Lundvall, S.; Barker, D. Dancing as searching with Deleuze—A study of what students in physical education teacher education express and experience in creative dance lessons. *Res. Danc. Educ.* **2022**, 1–16. [CrossRef]
47. Chan, C.H.; Ha, A.S.; Lander, N.; Ng, J.Y. Understanding the Teaching and Learning of Fundamental Movement Skills in the Primary Physical Education Setting: A Qualitative Study. *J. Teach. Phys. Educ.* **2023**, *42*, 621–630. [CrossRef]
48. Teng, M.F.; Wu, J.G. An Investigation of Learners' Perceived Progress during Online Education: Do Self-Efficacy Belief, Language Learning Motivation, and Metacognitive Strategies Matter? *Asia-Pac. Educ. Res.* **2024**, *33*, 283–295. [CrossRef]
49. Bao, L.; Soh, K.G.; Mohd Nasiruddin, N.J.; Xie, H.; Zhang, J. Unveiling the Impact of Metacognition on Academic Achievement in Physical Education and Activity Settings: A Comprehensive Systematic Review and Meta-Analysis of Qualitative Insights. *Psychol. Res. Behav. Manag.* **2024**, *17*, 973–987. [CrossRef]
50. Culajara, C.J. Enhancing Students' Learning Performance in Physical Education Through the Use of a Mobile Application with Game-Based Learning Approach. *Phys. Educ. Sports* **2023**, *2*, 1–9. [CrossRef]
51. Kumalasari, D.; Purwanta, H.; Aw, S. Comparative Analysis of Generation Z's Digital History Literacy in History Education Majors on Java Island: A Study of History Digital Literacy. *J. Educ. e-Learn. Res.* **2024**, *11*, 90–96. [CrossRef]

sustainability

MDPI

Article

Higher Education in China during the Pandemic: Analyzing Online Self-Learning Motivation Using Bayesian Networks

Jiang Li [1,2], Yating Chang [1,2], Shaobo Liu [1,2,*], Chang Cai [3,*], Qingping Zhou [3], Xiaoxi Cai [4], Wenbo Lai [5], Jialing Qi [1,2], Yifeng Ji [1,2] and Yudan Liu [1,2]

1 School of Architecture and Art, Central South University, Changsha 410083, China; lijiang@csu.edu.cn (J.L.); yatingchang@csu.edu.cn (Y.C.); jialingqi@csu.edu.cn (J.Q.); jyf0519@csu.edu.cn (Y.J.); ydt123@csu.edu.cn (Y.L.)
2 Human Settlements Research Center, Central South University, Changsha 410083, China
3 School of Mathematics and Statistics, Central South University, Changsha 410083, China; qpzhou@csu.edu.cn
4 School of Art and Design, Hunan First Normal University, Changsha 410205, China; xiaoxi@hnu.edu.cn
5 School of Architecture, South China University of Technology, Guangzhou 510006, China; lailaiwb@scut.edu.cn
* Correspondence: liushaobo@csu.edu.cn (S.L.); 212111095@csu.edu.cn (C.C.)

Abstract: The COVID-19 pandemic has led to an unprecedented shift towards online learning, compelling university students worldwide to engage in self-directed learning within remote environments. Despite the increasing importance of online education, the factors driving students' motivation for self-directed online learning, particularly those involving economic incentives, have not been thoroughly explored. This study aims to address this gap by analyzing large-scale data collected from 19,023 university students across China during the pandemic. Using mixed Bayesian networks and multigroup structural equation modeling, the study explores the complex relationships between personal characteristics, academic characteristics, the academic environment, and students' motivation for self-directed online learning. The results reveal significant associations between online self-directed learning motivation and personal characteristics. such as gender and age, academic characteristics, such as education level and learning incentives, and the geographic location of the school within the academic environment. Moreover, the causal relationship between school location and online self-directed learning motivation varies by gender and educational level. This research not only provides new empirical support for the theoretical framework of online learning motivation but also contributes to the broader fields of educational psychology and online learning research.

Keywords: online self-directed learning motivation; online search; learning-incentive money; hybrid Bayesian networks; multigroup structural equation modeling

Citation: Li, J.; Chang, Y.; Liu, S.; Cai, C.; Zhou, Q.; Cai, X.; Lai, W.; Qi, J.; Ji, Y.; Liu, Y. Higher Education in China during the Pandemic: Analyzing Online Self-Learning Motivation Using Bayesian Networks. *Sustainability* **2024**, *16*, 7330. https://doi.org/10.3390/su16177330

Academic Editors: Savvas A. Chatzichristofis, Jian-Hong Ye, Yung-Wei Hao and Yu-Feng Wu

Received: 22 July 2024
Revised: 14 August 2024
Accepted: 19 August 2024
Published: 26 August 2024

1. Introduction

In early 2020, the outbreak of the COVID-19 pandemic led to the largest educational disruption in history [1], and several countries and regions worldwide took measures to suspend classes in response to the outbreak. The effects of the outbreak did not spare extracurricular educational institutions in China; as universities went into lockdown, students could not participate in offline extracurricular training. Thus, after-school providers had to shift from traditional teaching models to online ones to reduce direct person-to-person contact [2–4]. In this context, the transition to online learning in higher education has accelerated, prompting a large number of university students to engage in self-directed learning within remote environments [5]. However, compared to primary and secondary school students, university students face the dual challenges of rapidly updating knowledge and increased employment pressures. To be able to keep up with the times and enhance their competitiveness, they usually take the initiative to study to acquire increasingly advanced knowledge and skills.

Proactive learning is not merely about meeting employment challenges, but it is also about achieving personal psychological fulfilment and self-worth. However, maintaining university students' learning motivation, particularly in the context of online courses during the pandemic, has become a major challenge. Research has shown that different types of social support, professional courses, learning institutions, learning environments, and other factors affect self-directed learning motivation [6–10]. For example, Baeten et al. [6] studied the effects of different learning environments on students' motivation to learn and achieve. Li et al. [2] found that students' grade level, gender, and learning environment may be important factors affecting academic performance in online education, and Wei et al. [11] found that groups with self-directed learning motivation showed higher perceived learning outcomes than their peers.

While these studies have significantly contributed to understanding learning motivation in traditional educational settings, the factors driving university students' self-directed online learning, especially those involving economic incentives, have not been sufficiently explored. Economic incentives, such as learning-incentive money (which may include rewards, subsidies, or scholarships), are particularly relevant in online learning environments. These financial incentives can motivate students by providing tangible rewards for academic achievements, which, in turn, can enhance their self-directed learning motivation. For example, monetary incentives have been shown to increase students' engagement and effort in laboratory-based tasks, but their effectiveness in natural educational settings has yielded mixed results [12,13]. Understanding how learning-incentive money interacts with other personal and academic factors to influence motivation is crucial for developing effective strategies in online education.

Most research to date has focused on small sample sizes and has primarily analyzed offline learning models, making it difficult to generalize findings to the broader context of the global shift to online learning. Therefore, it is necessary to analyze data from large samples of university students during the COVID-19 pandemic to explore how personal characteristics, academic characteristics, and the academic environment influence online self-directed learning motivation. Additionally, conducting subgroup analyses among undergraduate, master's, and doctoral students can provide deeper insights into how these variables affect different groups.

Based on this, the objectives of this study are: (1) to analyze how university students' personal characteristics, academic characteristics, and academic environment variables influence online self-directed learning motivation using large-scale sample data, and to explore the causal relationships between these factors; (2) to conduct subgroup analyses of undergraduates, master's students, and doctoral candidates to gain a deeper understanding of how these variables influence online self-directed learning motivation across different groups. The findings of this study will not only provide new empirical support for the theoretical framework of online learning motivation but also contribute to the fields of educational psychology and online learning research. Furthermore, these insights will offer valuable information and recommendations for the development of future online education platforms, course design, and teaching strategies, particularly in promoting the sustainable development of online education in a globalized context.

2. Theoretical Framework

2.1. Self-Determination Theory and Online Learning Motivation

Self-Determination Theory (SDT), developed by Ryan and Deci [14], is a widely recognized framework for understanding motivation within educational settings. SDT categorizes motivation into two types: intrinsic and extrinsic. Intrinsic motivation arises from the inherent enjoyment and interest in the activity itself, while extrinsic motivation is driven by external rewards or pressures. According to SDT, three basic psychological needs—autonomy, competence, and relatedness—must be satisfied to foster intrinsic motivation and optimal learning performance.

In the context of online autonomous learning, these needs are particularly critical. Autonomy refers to a student's sense of control over their learning process, while competence refers to their confidence and ability to successfully complete online learning tasks. Research indicates that students with intrinsic or self-determined motivation typically achieve more positive learning outcomes [15,16]; for example, they are better able to adapt to academic life, have lower levels of perceived stress, and can engage in more sustained learning [17,18]. With the rapid advancement of technology, the manner in which people learn or acquire new knowledge has changed, and one major change is that an increasing number of students are choosing to receive education online [19]. Autonomy is a key feature in online learning [20,21]. This surge in demand for online learning resources further underscores the importance of self-directed learning motivation in this new educational environment.

2.2. The Impact of Personal Characteristics on Online Learning Motivation

Students' personal characteristics, such as gender, age, and educational level, are closely related to their learning motivation [22]. A study found significant gender differences in self-directed learning motivation [20], where girls usually show a higher autonomous motivation to learn [23]; this trend may contribute to their higher academic achievement. Ethnicity is associated with student autonomy and, during distance learning throughout the pandemic, curiosity among Latino students enrolled in ethnic studies courses improved during the school year, although students' stress and motivation were lower. Additionally, research has indicated that, as people age, they tend to develop a stronger sense of autonomy; older individuals are better at internalizing their goals and personal initiatives than younger individuals.

2.3. The Role of Academic Characteristics in the Formation of Learning Motivation

Academic characteristics, such as school atmosphere and reward systems [24], play distinct roles in influencing students' motivation and online learning behaviors. Economic Incentive Theory posits that monetary rewards can significantly influence behavior by increasing the perceived benefits of specific actions [25,26]. In educational settings, economic incentives have been shown to enhance student engagement, particularly in environments where motivation may be low. In online learning, where immediate guidance and timely feedback are often challenging to provide, learning-incentive funds can serve as an effective external motivator, enhancing students' extrinsic motivation and thereby promoting greater engagement in their studies [13]. It is worth noting that reward mechanisms in the form of external incentives may also lead to learner overdependence, with various implications for human behavior [27]. In addition, a correlation exists between academic climate and academic performance. Negative school belonging has been noted to negatively impact intrinsic motivation and perceived learning [28]. Higher-level schools offer a more comprehensive curriculum, better-quality educational resources, and a learning environment suitable for the development of self-directed learning skills. Such schools tend to encourage active inquiry and self-directed learning processes, and they may provide more external resources and opportunities to promote the development of self-learning skills.

2.4. The Impact of Academic Environment on Learning Motivation

The academic environment, including school type, geographic location, and economic background, often influences the resources and opportunities available for self-directed learning. Schools' geographic location and economic level typically correlate with access to richer online learning resources. In areas of higher economic development, schools are likely to have more resources available for pedagogical innovation, thus providing students with a wealth of learning material and environments that enhance their self-directed learning. Conversely, in less-affluent economies, the investment capabilities of schools and families in education may be constrained, potentially inhibiting the promotion of self-directed learning. Students from lower socioeconomic backgrounds and rural communities may require additional support to overcome educational challenges—particularly those exacerbated by the COVID-19 pandemic. Furthermore, the nature of a school (i.e., whether it is academic, technical–vocational, or comprehensive) influences the development of self-learning capabilities. Academic-oriented schools may place greater emphasis on theoretical knowledge and research skill development, thereby facilitating self-directed learning, whereas technical vocational schools may prioritize skills-based and practical learning experiences. Overall, the evolution of self-learning is affected by multiple factors: a school's type, nature, and geographic location can directly or indirectly impact the distribution of educational resources, selection of educational strategies, and improvement of educational quality, all of which, in turn, shape students' capacity for self-directed learning.

This study integrates principles from Self-Determination Theory and Economic Incentive Theory to develop a comprehensive model that explains the factors influencing college students' motivation for self-directed online learning. The model posits the following:

(1) Learning-incentive funds, as an external motivating factor, can enhance extrinsic motivation while potentially fostering self-directed learning motivation in an online setting.
(2) Personal characteristics (such as gender and age), academic characteristics (such as educational background and reward systems), and the academic environment (such as school type and geographic location) significantly impact students' learning motivation, with these factors being particularly crucial in online learning environments.
(3) The interaction among these variables is expected to reveal how different groups respond to the challenges and rewards of online learning.

3. Method

3.1. Data Collection

In our study, we used the online search frequency of students on the learning platform as a key indicator of online self-directed learning motivation. Online search frequency refers to the rate at which students actively search for learning resources and information through the platform during their studies. This behavior reflects students' willingness and ability to actively acquire knowledge, making it an external manifestation of self-directed learning motivation [16,29]. Furthermore, online search frequency data is easily accessible and can provide real-time insights into students' learning motivation, making it a feasible and meaningful indicator for large-scale studies.

The survey data collected during the pandemic was provided by the Chinese educational technology company Zhuolang Technology. This company offers a wide range of content across various academic disciplines, enabling both undergraduate and graduate students to easily access academic materials and learning resources. Additionally, the company provides learning incentives based on students' academic performance, including rewards, subsidies, or scholarships. These financial incentives can be used by students to purchase learning materials, pay for course fees, or serve as rewards for completing specific learning tasks.

The raw data were collected in the form of web search logs, which detailed students' search activities on the platform. Each log entry included information on search frequency, timestamps, and the types of content accessed. Additionally, demographic and academic information about the students, such as gender, age, ethnicity, educational level, and school characteristics, was included. To ensure the quality and accuracy of the data, we employed several data cleaning procedures. First, we removed any duplicate or incomplete entries from the dataset. Next, we standardized the format of the remaining data to ensure consistency across different variables. For categorical variables, such as search frequency, we categorized the data into five ordered levels based on the actual number of searches. Finally, we identified and appropriately handled any outliers that could impact the analysis, maintaining the integrity of the dataset. As a result, these data are highly reliable and representative, providing a solid foundation for analyzing and revealing differences and similarities between groups.

We utilized hybrid Bayesian networks (BN) and multigroup structural equation modeling (SEM) for the causal analysis of the data (Figure 1). In the first stage, we considered gender, age, and ethnicity as students' personal characteristics, and school level, education level, length of time spent in school, and learning-incentive money as their academic characteristics. The number of learning incentives on the online education platform is the number of incentives students received on that platform. We used school type, school nature, school economic geographic location, and school regional geographic location as academic environment variables. We adopted a hybrid BN-based approach to identify causal links between online self-directed learning motivation and learner characteristics (student personal characteristics, academic characteristics, and academic environment variables) among university students. In the second stage, we extracted online self-directed learning motivation as the core key variable and built a multigroup SEM with high causal arc strength as a premise from which important causal relationships were mined. In the third stage, we used multigroup SEM for group difference analysis to determine whether significant differences existed between the groups.

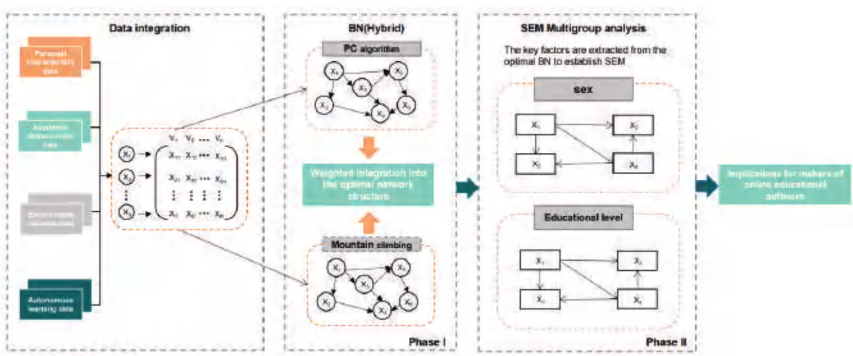


Figure 1. Flowchart.

3.2. Bayesian Networks

The BN is a probabilistic graphical model consisting of a set of random variables and a Directed Acyclic Graph (DAG) [30], where nodes denote random variables and edges denote dependencies between these variables. This model allows the dependencies encoded in the DAG to be mapped to the probability space through conditional independence relations, thus performing quantitative probabilistic inference under uncertainty. BN can effectively handle large sample data to determine causal relationships among variables related to online self-directed learning motivation. BN is capable of dynamically adjusting its network based on the data presented or input, enabling it to successfully manage large and multivariate datasets, as well as handle incomplete or uncertain knowledge or information [31,32]. However, BN is less robust in theoretical interpretation compared to SEM [33,34]. Therefore, some studies combine these two models to establish causal relationships [35,36].

3.2.1. Dependency Establishment

Let $U = (x_1, x_2, \ldots, x_{12})$ be the set of input points and a total of one variable which represent SEX, MON, FRE, LEN, EDU_L, AGE, SCH_L, NAT, ECO_L, LOC, TYP, SRN shown in Table 1. A BN on a set of variables U is a network structure which is a DAG on U and a set of probability tables [37]:

$$B_P = \{P(x_i | pa(x_i), x_i \in U)\} \quad (1)$$

where $pa(x_i)$ is the set of the antecedents of x_i in BN and $i = 1, 2, 3, \ldots, 12$.

A BN represents joint probability distributions (2):

$$P(U) = \prod_{x_i \in U} p(x_i | pa(x_i)) \quad (2)$$

3.2.2. Structure Learning for Weighted Hybrid Peter-Clark–Hill-Climbing (PC-HC) Algorithm

The Peter-Clark–Hill-Climbing (PC–HC) algorithm, based on a weighted mixture, takes full advantage of the independence of scores and conditions. Among them, the method based on the conditional independence test is more efficient and can obtain the optimal global solution [38]. The scoring algorithm is a local search strategy with the problem of getting caught in local poles and saddle points, but it is more flexible in the way it changes the network's structure [39]. The hybrid algorithm first obtains the graph structure according to the Peter-Clark (PC) algorithm and then uses the Hill-Climbing (HC) algorithm to connect the learned structures of each subgraph, and weighted corrections are made to the two graph structures to obtain the optimal BN structure.

① Peter-Clark algorithm (PC)

As an algorithm based on the assumption of conditional independence, the PC algorithm can be flexibly combined with various conditional independence tests to infer causality between variables and thus obtain the optimal global solution [40]. The core idea of the PC algorithm is conditional independence using the following d-separation principle: the set of nodes I can separate nodes m and n, and no valid path exists between m and n if and only if I is given.

② Hill-Climbing algorithm (HC)

As a score-based structure learning algorithm, the HC algorithm usually starts with an empty graph. It explores the search space of the graph by adding, deleting edges, and edges to maximize the target score and stops iterating when changing the structure of the edge neighbors; then, the score no longer increases to obtain the optimal structure of the graph, which may also be a local optimum [41,42].

After obtaining network structures G1 and G2 based on the two methods, the new network structure G is obtained by mixing and weighting the networks using the set weight vectors and selecting the substructures for subsequent analysis.

3.3. Multigroup Structural Equation Model (SEM)

SEM is a causal modeling approach that combines causal information with statistical data to quantitatively assess the relationship between the variables under study [43]. It allows for better control of measurement errors than general regression analysis and supports the construction of complex multivariate models. The main purpose of multigroup SEM analysis test is to assess whether the path model map can be adapted for other different groups when matched to one group [44], as defined in [45].

$$\eta = B\eta + \Gamma\xi + \zeta \quad (3)$$

where $\eta = m \times 1$ is a vector of latent variables, in this project, $m = 6$ is reflected in the key network which supports the implementation of the structural equation model; $\xi = n \times 1$ is a vector of latent exogenous variables; $B = m \times m$ is a matrix of coefficients associated with latent endogenous variables; $\Gamma = m \times n$ is a matrix of coefficients associated with latent exogenous variables; and $\zeta = m \times 1$ is a vector of error terms associated with endogenous variables.

Table 1. Descriptive analysis (n = 19,023).

Variables	Abbreviation	Count	Percentage (%)
Gender	sex		
Men		12,333	0.65
Women		6690	0.35
Ethnicity	ETH		
Ethnic Han		17,368	0.91
Ethnic Minorities		1655	0.09
Age	AGE		
≤18		248	0.01
19		2686	0.14
20		4843	0.25
21		4464	0.23
22		3143	0.17
≥23		3693	0.19
School level	Sch_l		
First-rate universities and disciplines		2979	0.16
Full-time college		11,535	0.60
Higher vocational college		4509	0.24
School-running nature	SRN		
Public school		15,474	0.81
Private school		2630	0.14
Independent college		905	0.05
Chinese–foreign cooperatively run schools		14	0.00
Length of schooling (year)	LEN		
1.5		3	0.00
2		838	0.04
2.5		68	0.00
3		5555	0.29
3.5		1	0.00
4		11,806	0.62
5		729	0.05
6		2	0.00
8		21	0.00
Type of school	typ		
Research university		1432	0.08
Applied university		13,061	0.68
Technical university		4530	0.24
Education level	Edu_l		
Junior college		5092	0.27
Undergraduate		13,050	0.69
Master's degree		797	0.04
doctor		84	0.00
Online search frequency	fre		
Low frequency		4400	0.23
Less frequency		3145	0.17
Average frequency		3859	0.20
High frequency		3806	0.20
Very high frequency		3814	0.20
School area geographic location	loc		
South China		2352	0.12
North China		2886	0.15
East China		4816	0.25
Southwest		2238	0.12
Northwest		1451	0.08
Central China		3434	0.18
Northeast China		1846	0.10

Table 1. *Cont.*

Variables	Abbreviation	Count	Percentage (%)
Economic and geographic location of the school	Eco_l		
First-tier city		7809	0.41
Second-tier city		4978	0.26
Third-tier city		3302	0.17
Fourth-tier city		2006	0.11
Fifth-tier city		928	0.05
Money (CNY)	mon		
≤5		111	0.01
5–10		11,172	0.58
10–15		206	0.01
15–20		7191	0.38
≥20		343	0.02

4. Results

This section presents the main results of the study, with a detailed focus on how personal characteristics, academic factors, and academic environment variables influence college students' motivation for online self-study. By analyzing large-scale sample data, we identify key relationships between these variables and online learning motivation, and examine the differences across various groups (e.g., gender, education level). These findings will serve as the foundation for the subsequent discussion, where we will further explore how these results address the research questions posed in this paper.

4.1. Descriptive Analysis

A total of 19,023 items of data for 2392 Chinese universities were obtained from the survey. According to the results of the descriptive statistical analysis (Table 1), men accounted for 65% of the sample size, and women accounted for 35%—slightly less than men. In terms of age distribution, students in their 20s predominated, accounting for 25% of the sample. In terms of school level, students from top-ranking and world-class universities accounted for 16% of the sample size; students from general universities accounted for 60%, and students from higher education institutions for 24%. In terms of the nature of schooling, most students were from public schools, accounting for 81% of the sample. In terms of student duration, most students (42% of the sample) were in four-year programs. In terms of the type of school, applied universities were the majority, accounting for 68% of the sample, followed by technical universities at 24% and research universities at 8%. In terms of students' education level, undergraduate degrees were predominant (69% of the sample). In terms of the frequency of searches, this was very high for 20% of the sample and very infrequent for 23%. In terms of the schools' regional geography, students were more evenly distributed, with 25% of the sample located in East China. In terms of the schools' economic geography, most students were in first-tier cities, accounting for 41% of the sample.

4.2. Results of BN

In this study, a hybrid BN was adopted for structural learning to obtain causal links between variables, and the data were preprocessed using R 4.2.0 to obtain DAGs and key networks, as shown in Figure 2.

The results obtained using the PC and HC algorithms were generally consistent. However, to enhance the interpretability of the structure and make the results easier to understand and accept, we adjusted the network structure according to a priori knowledge; that is, we used a weighted integration method to reconstruct a more realistic hybrid BN (Figure 2b). Figure 2a is found to be more credible than Figure 2b based on domain knowledge, setting the percentage of network credibility of Figure 2a to 0.7.

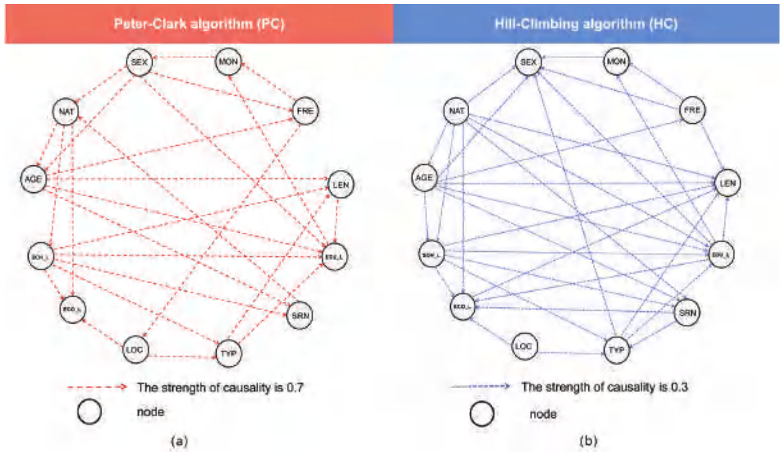


Figure 2. (**a**) An initial BN model using the PC algorithm; (**b**) an initial BN model using the HC algorithm.

Figure 3 shows the optimal BN for this study, where arcs with a scale close to 1 are considered more robust in terms of causality, whereas the opposite is true for arcs close to 0. We extracted substructures using online self-learning dynamics as key variables to provide a basis for subsequent group difference analysis and obtained the following key network using dual core search frequency and learning-incentive money as leaf or parent nodes (Figure 3). The independent learning variables are mainly causally related to gender and age in the students' personal characteristics, education level in the students' academic characteristics, and the school area's geographic location in the academic environment variables. In contrast, in this study, the independent learning variables were not related to ethnicity, school level, year, school type, nature of school operation, and economic geographic location of the school.

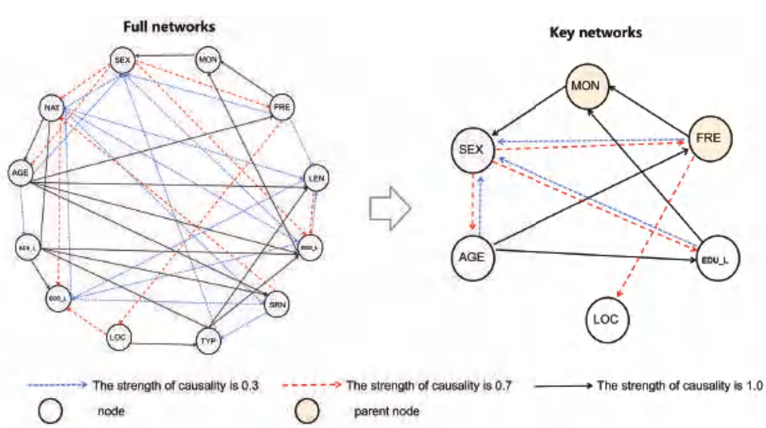


Figure 3. Key Networks.

4.3. Results of SEM

After the BN combined a causal model that gave the maximum likelihood based on the data, we used SEM to check which nodes and arcs were reasonable and determine the robustness of the BN's bootstrap approach. In addition, to enhance the interpretability of the structure and make the results more understandable and acceptable, we slightly adapted the initial structure to the domain knowledge.

Table 2 presents the final model fit indices. The root mean squared error of approximation (RMSEA) [46] is less than or equal to 0.08; the incremental fit index (IFI) is greater than or equal to 0.90; the comparative fit index (CFI) [47] is greater than or equal to 0.90, and the Goodness of Fit Index (GFI) is greater than or equal to 0.90, which is considered to be a good model fit [48]. The important fit indicators in this study are within the acceptable range of suggested values; therefore, the SEM model fits well.

Table 2. The fitting index value of multigroup structural equation modeling (n = 19,023).

	X^2	X^2/df	RMSEA	IFI	CFI	GFI
No-group measured value	66.151	9.450	0.021	0.951	0.950	0.999
Gender						
Women	5.34	0.995	0.025	0.972	0.972	0.998
Men	14.999	3	0.013	0.978	0.978	1
Education level						
Junior college	9.253	0.313	0.016	0.972	0.971	0.999
Undergraduate	19.799	4.95	0.017	0.971	0.971	0.999
Master's degree	9.968	2.492	0.041	0.524	0.088	0.995

Figure 4 shows the standardized path coefficients and significance levels between the variables in the study model. The SEM results largely confirmed the significant arc of BN, except that student search frequency was negatively correlated with school area and geographic location. In addition, the path coefficients for the total sample and the subgroups converged broadly, demonstrating stable causal pointing between the variables. Of these, in terms of motivation to learn independently online, online search frequency (β = 0.082, p = 0.001) significantly influenced learning-incentive money. Gender (β = 0.038, p = 0.001) and age (β = 0.135, p = 0.001) also significantly influenced the frequency of online searches.

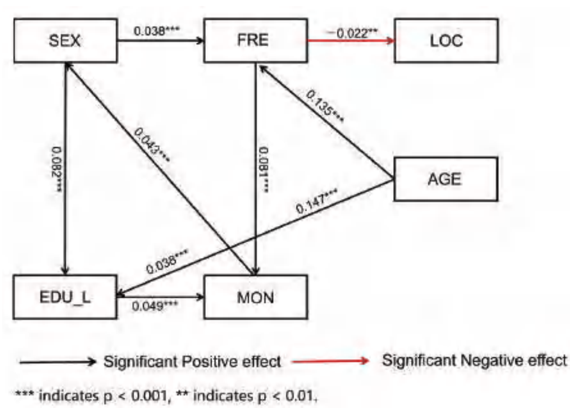


Figure 4. Diagram of the SEM.

4.4. Results of Multi-Group SEM

To determine whether the associations between subjective learning motivation and college students' age and geographic location of the school area differed across gender and education level, we used two variables—college students' gender and education level—as the basis for multigroup analysis to divide the model. The results (Table 2) showed that the important fit indicators were within the acceptable recommended values; therefore, the model fit was good.

Figure 5a shows the results of the multigroup SEM analysis based on gender. The SEM for women (Figure 5a) shows a positive and significant correlation between age and students' search frequency (β = 0.128, p = 0.001) and education level (β = 0.291,

p = 0.001), a positive and significant correlation between students' search frequency and learning-incentive money (β = 0.085, p = 0.001), and no correlation between students' search frequency and school area's geographic location (β = −0.016, p = 0.178); moreover, education level (β = 0.042, p = 0.001) was positively and significantly associated with learning-incentive money.

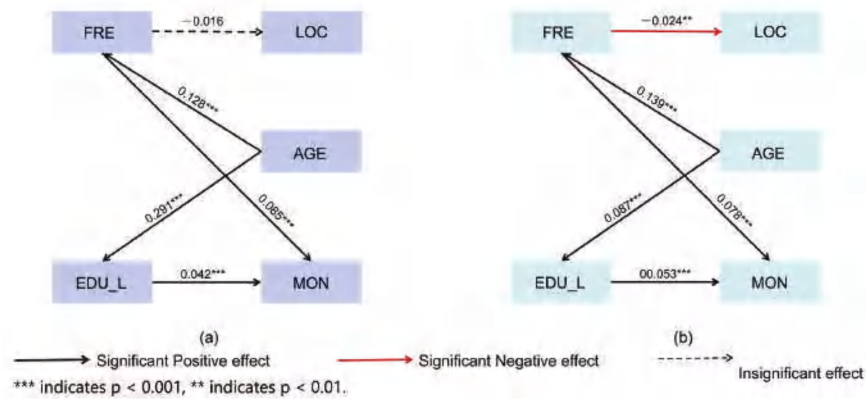


Figure 5. The results of multigroup SEM analysis based on gender: (**a**) women; (**b**) men.

The SEM for men (Figure 5b) showed a positive and significant correlation between age and students' search frequency (β = 0.139, p = 0.001) and education level (β = 0.087, p = 0.001), a positive and significant correlation between students' search frequency and learning-incentive money (β = 0.078, p = 0.001), and a negative and significant correlation between students' search frequency and school area geographic location (β = −0.024, p = 0.008); moreover, education level (β = 0.053, p = 0.001) was positively and significantly associated with learning-incentive money.

Then, we analyzed multi-cluster structural equation modeling based on education level. Figure 6 (right) shows the results of the multigroup SEM analysis based on education level. The model plots for students in junior college education and the individual path coefficients (Figure 6a) show a negative correlation between students' search frequency and school area's geographic location (β = −0.033, p = 0.020), a negative correlation between genders (β = 0.036, p = 0.010), a positive correlation between age (β = 0.127, p = 0.001) and students' search frequency, and a positive and significant correlation between students' search frequency and learning-incentive money (β = 0.080, p = 0.001).

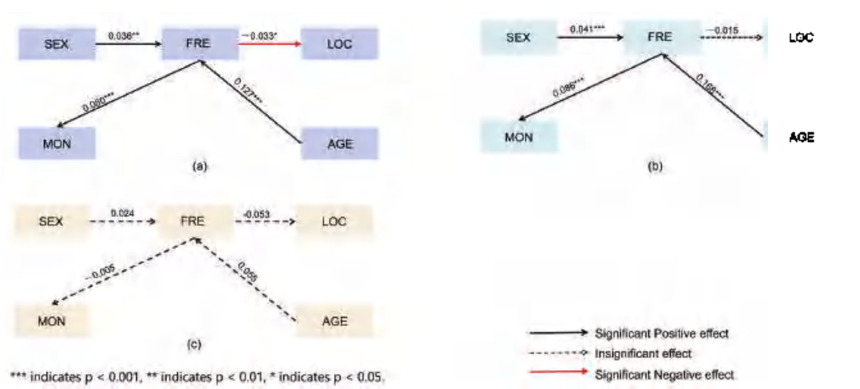


Figure 6. The results of multigroup SEM analysis based on education level: (**a**) tertiary education; (**b**) bachelor's degree; (**c**) postgraduate qualifications.

The model plots for undergraduate students and the individual path coefficients (Figure 6b) show that students' search frequency is not related to the school area's geographic location (β = −0.015, p = 0.083); gender (β = 0.041, p = 0.001) and age (β = 0.168, p = 0.001) are positively related to students' search frequency; and students' search frequency is positively and significantly related to learning-incentive money (β = 0.086, p = 0.001).

The model plots for students with postgraduate qualifications and the individual path coefficients (Figure 6c) show that students' search frequency was independent of school area geographic location (β = −0.053, p = 0.117); gender (β = 0.024, p = 0.476) and age (β = 0.055, p = 0.101) were independent of students' search frequency; and students' search frequency was independent of learning-incentive money (β = −0.005, p = 0.874).

The findings for the postgraduate group are contrary to those of the overall relationship graph. Potential influences may cause the relationships for the variables explored to not hold, and this group may be more oriented toward academic studies than toward traditional exam education. Therefore, the learning characteristics explored did not have a significant causal relationship in this group.

5. Discussion

In terms of students' personal characteristics, gender and age were significant factors influencing motivation for autonomous learning. Furthermore, age had a slightly higher impact on self-directed learning in the men's group than in the women's group, which may be due to the traditional division of labor between men and women, where men are primarily responsible for providing for their families. Although the number of women has risen in recent years, and their contribution to household income has grown, overall, the core of household income remains male. Under the same social pressures, older men, who are often pressured to be employed to support their families, may be more anxious about learning independently (and more motivated to do so) than women. In addition, the effect of age on educational attainment is significantly more relevant in the men's cohort than in the women's cohort, which may be due to the fact that age plays a crucial role in women's post-college human capital investment decisions. Zhang et al. note that, according to Census and CFPS data, women who enter school at a relatively later age are less likely than men to enter graduate school [49]. The act of delaying employment and marriage as they age puts women at a disadvantage in the labor and marriage markets, which is considered an opportunity cost of investing in human capital after college.

In terms of students' academic characteristics, learning-incentive money was positively correlated with motivation for autonomous learning. This is consistent with previous research showing that small rewards enhance autonomous motivation [50]. Learning-incentive money is a moral and material incentive to help students realize their self-worth, which motivates college students to study independently online and prompts them to engage more actively in their own behavior. Appropriate rewards can replace feedback from online educators, making it possible to achieve desired learning outcomes in a learning environment with limited personal interaction. The education level was also significantly and positively associated with learning-incentive money; this may be because less-educated students have less awareness of their careers, lack initiative and self-motivation to learn, and therefore receive less learning-incentive money. In addition, students with specialized qualifications tend to have a pronounced herd mentality and generally less self-control compared to undergraduate students and are prone to indulging in these establishments, which are conveniently located and rich in recreational venues [51]. Therefore, a region's geographic location influences the self-learning motivation of students with tertiary qualifications. The doctoral degree group is more inclined to academic research than traditional examination-based education; therefore, in this group, we found no significant causal relationship among the learning characteristics explored.

In terms of academic environment, students' motivation for self-learning was often affected by the geographic location of the school area. For example, the level of Internet development in China's eastern coastal regions was significantly higher than that in the central and western regions, and the level of network construction in cities was significantly higher than that in rural areas. The existence of poorer network environments in remote areas, the lag phenomenon when learning online, and the inability to operate and experience technological software all directly affected the students' learning progress and search results. We also found that the women's and men's groups diverged in terms of the impact of the school area's geographic location on self-directed learning, and this gender difference may be related to cultural and social expectations. According to Ong, cultural, social, and institutional biases contribute to women's inability to fit into the current culture's power spaces, which are more pronounced in science, technology, engineering, and mathematics (STEM) professions [52]. Some cultures tend to place more emphasis on men's roles in the workplace and society, whereas women are more often considered the primary caregivers of the family. This cultural and societal expectation may lead men to focus more on advantages of geographic location, such as school reputation and ranking.

6. Conclusions

University students' education is essential for the dissemination of knowledge, technological innovation, and sustainable social development. Although the COVID-19 pandemic has transformed traditional offline classroom education into online education, higher education still faces new challenges and opportunities. Taking Chinese college students as the research object, we selected factor variables related to their online learning, and models and analyzed their motivation for online self-directed learning by means of hybrid BN and multigroup SEM. The main findings and conclusions are summarized as follows.

Overall, the SEM test results were broadly consistent with those of the mixed BN key network and multiple SEM subgroups, suggesting a stable causal relationship between the factors. Self-learning motivation was mainly related to gender and age in students' personal characteristics, to education level in students' academic characteristics, and to the school area's geographic location in the academic environment variables. Learning-incentive money had a significant positive effect on self-directed learning motivation for different gender groups; more specifically, it was not related to the school area's geographic location in the women's group, whereas it had a negative correlation with it in the men's group. Among groups with different education levels, the self-learning motivation of students with specialized degrees had a negative correlation with the school area's geographic location, whereas the online search frequency of students with bachelor's degrees was not related to the geographic location of their school area. These findings help to improve the understanding of college students' motivation for online self-directed learning and recognize the future role of online educational environments in promoting health. Based on the study's findings, future online education platforms can offer personalized learning incentives tailored to students' gender, age, and educational background, thereby enhancing their self-directed learning motivation. This study not only expands the current understanding of how demographic characteristics, economic incentives, and educational factors influence learning motivation, providing a foundation for further research in educational psychology and online learning, but also contributes to the sustainable development of online education in a globalized context.

7. Limitations and Future Prospects

The study identified several key factors that influence students' self-directed learning motivation in online environments, including gender, age, educational background, and the impact of various incentives. These findings offer valuable insights for the design and development of future online education platforms.

7.1. Personalized Learning Incentives

Based on our research results, future online education platforms can implement personalized learning incentives tailored to the specific characteristics of students. For example, gender-specific motivational strategies can be developed to address the unique learning preferences and challenges faced by male and female students. This is particularly important given the study's findings that gender plays a significant role in self-directed learning motivation. Similarly, age-appropriate rewards can be designed to align with the developmental stages and cognitive needs of different age groups. Additionally, rewards suited to students' educational backgrounds can be created to support varying levels of academic preparedness and experience, thereby enhancing self-directed learning motivation. By adopting these targeted approaches, online education platforms can more effectively engage students and foster a more productive learning environment.

7.2. Tiered Learning Experiences

The study also highlights the different motivational needs across educational levels (undergraduate, master's, doctoral). Considering the volume of learning and research tasks at each stage, course designers and instructors are advised to implement tiered online learning experiences for undergraduates. For example, a virtual research community could be created for master's students, including additional learning sections on research submissions and writing, serving as a reliable source for academic support, enabling students to receive timely feedback from their peers.

7.3. Comprehensive Support Systems

Recognizing the importance of various academic characteristics and environmental factors identified in the study, system developers can create multiple channels to facilitate students' help-seeking behavior, such as online Q and A forums or collaborative help-seeking tools embedded in learning management systems [49,51]. Platforms can offer different types of rewards, such as incentives for users to reopen the app, significantly increasing the likelihood of repeated use. These platforms could combine rewards, like points, levels, badges, tasks, and leaderboards, with visual feedback to encourage ongoing participation and enhance students' motivation for online self-directed learning [52].

7.4. Limitations

The ability to learn independently online is influenced by many other factors, such as different online search models, teachers' teaching methods, the online platform's management style, students' personality and behavior, and their family's educational background. Furthermore, as the participants in this study were drawn from various disciplines, the results derived can only provide general insights into the online learning profiles of university students. Students from specific fields or domains are likely to exhibit learning tendencies that differ from those identified in this study and necessitate further research. Therefore, in the future, we intend to extend this analysis to include additional data from other large online-learning courses across various disciplines. In addition, other influencing factors, such as learner time interactions, different online search patterns, teachers' teaching-method preferences, and family educational background, will be utilized to present a relatively comprehensive set of models for detailed analysis.

Author Contributions: Conceptualization, Y.C.; methodology, C.C. and Q.Z.; formal analysis and investigation, Y.C., C.C., J.Q. and Y.L.; resources, J.L. and X.C.; writing—original draft preparation, Y.C.; writing—review and editing, J.L. and Y.J.; supervision, J.L. and S.L.; funding acquisition, S.L. and W.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (grant numbers 52108049, 42177072, and 51909283), the Natural Science Foundation of Hunan Province, China (grant numbers 2020JJ4711 and 2023JJ30182), the Humanities and Social Science Fund of Ministry of Education of China (grant number 20YJCZH003), the Key Research and Development Program of Hunan Province in China (grant number 2020WK2001), and the Fundamental Research Funds for the Central Universities of Central South University (grant number 2021XQLH116).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Argyropoulos, K.V.; Serrano, A.; Hu, J.; Black, M.; Feng, X.; Shen, G.; Call, M.; Kim, M.J.; Lytle, A.; Belovarac, B.; et al. Association of Initial Viral Load in Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Patients with Outcome and Symptoms. *Am. J. Pathol.* **2020**, *190*, 1881–1887. [CrossRef]
2. Li, J.; Che, W. Challenges and coping strategies of online learning for college students in the context of COVID-19: A survey of Chinese universities. *Sustain. Cities Soc.* **2022**, *83*, 103958. [CrossRef] [PubMed]
3. Segbenya, M.; Bervell, B.; Minadzi, V.M.; Somuah, B.A. Modelling the perspectives of distance education students towards online learning during COVID-19 pandemic. *Smart Learn. Environ.* **2022**, *9*, 13. [CrossRef]
4. van der Velde, M.; Sense, F.; Spijkers, R.; Meeter, M.; van Rijn, H. Lockdown Learning: Changes in Online Foreign-Language Study Activity and Performance of Dutch Secondary School Students During the COVID-19 Pandemic. *Front. Educ.* **2021**, *6*, 712987. [CrossRef]
5. Lomellini, A.; Lowenthal, P.R.; Snelson, C.; Trespalacios, J.H. Higher education leaders' perspectives of accessible and inclusive online learning. *Distance Educ.* **2022**, *43*, 574–595. [CrossRef]
6. Baeten, M.; Dochy, F.; Struyven, K. The effects of different learning environments on students' motivation for learning and their achievement. *Br. J. Educ. Psychol.* **2013**, *83*, 484–501. [CrossRef]
7. Büchele, S. Evaluating the link between attendance and performance in higher education: The role of classroom engagement dimensions. *Assess. Eval. High. Educ.* **2021**, *46*, 132–150. [CrossRef]
8. Kim, D.; Jung, E.; Yoon, M.; Chang, Y.; Park, S.; Kim, D.; Demir, F. Exploring the structural relationships between course design factors, learner commitment, self-directed learning, and intentions for further learning in a self-paced MOOC. *Comput. Educ.* **2021**, *166*, 104171. [CrossRef]
9. Paechter, M.; Maier, B.; Macher, D. Students' expectations of, and experiences in e-learning: Their relation to learning achievements and course satisfaction. *Comput. Educ.* **2010**, *54*, 222–229. [CrossRef]
10. Palaniappan, K.; Md Noor, N. Gamification Strategy to Support Self-Directed Learning in an Online Learning Environment. *Int. J. Emerg. Technol. Learn. (iJET)* **2022**, *17*, 104–116. [CrossRef]
11. Wei, X.; Saab, N.; Admiraal, W. Do learners share the same perceived learning outcomes in MOOCs? Identifying the role of motivation, perceived learning support, learning engagement, and self-regulated learning strategies. *Internet High. Educ.* **2023**, *56*, 100880. [CrossRef]
12. O'Neil, J.H.F.; Sugrue, B.; Baker, E.L. Effects of Motivational Interventions on the National Assessment of Educational Progress Mathematics Performance. *Educ. Assess.* **1995**, *3*, 135–157. [CrossRef]
13. Schwab, J.F.; Somerville, L.H. Raising the Stakes for Online Learning: Monetary Incentives Increase Performance in a Computer-Based Learning Task Under Certain Conditions. *Front. Psychol.* **2022**, *13*, 780301. [CrossRef] [PubMed]
14. Ryan, R.M.; Deci, E.L. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* **2000**, *55*, 68–78. [CrossRef] [PubMed]
15. Pan, Y.; Gauvain, M. The continuity of college students' autonomous learning motivation and its predictors: A three-year longitudinal study. *Learn. Individ. Differ.* **2012**, *22*, 92–99. [CrossRef]
16. Keller, J.M. First principles of motivation to learn and e3-learning. *Distance Educ.* **2008**, *29*, 175–185. [CrossRef]
17. Gottfried, A.E. Chapter Three—Academic Intrinsic Motivation: Theory, Assessment, and Longitudinal Research. In *Advances in Motivation Science*; Elliot, A.J., Ed.; Elsevier: Amsterdam, The Netherlands, 2019; Volume 6, pp. 71–109.
18. Lin, Y.G.; McKeachie, W.J.; Kim, Y.C. College student intrinsic and/or extrinsic motivation and learning. *Learn. Individ. Differ.* **2003**, *13*, 251–258. [CrossRef]
19. Xu, Z.; Zhao, Y.; Liew, J.; Zhou, X.; Kogut, A. Synthesizing research evidence on self-regulated learning and academic achievement in online and blended learning environments: A scoping review. *Educ. Res. Rev.* **2023**, *39*, 100510. [CrossRef]
20. Zhang, Z.; Maeda, Y.; Newby, T. Individual differences in preservice teachers' online self-regulated learning capacity: A multilevel analysis. *Comput. Educ.* **2023**, *207*, 104926. [CrossRef]
21. Chiu, T.K.F. Applying the self-determination theory (SDT) to explain student engagement in online learning during the COVID-19 pandemic. *J. Res. Technol. Educ.* **2022**, *54* (Suppl. 1), S14–S30. [CrossRef]
22. Steinmayr, R.; Lazarides, R.; Weidinger, A.F.; Christiansen, H. Teaching and learning during the first COVID-19 school lockdown: Realization and associations with parent-perceived students' academic outcomes. *Z. Für Pädagogische Psychol.* **2021**, *35*, 85–106. [CrossRef]
23. Vantieghem, W.; Van Houtte, M. Differences in Study Motivation Within and Between Genders: An Examination by Gender Typicality Among Early Adolescents. *Youth Soc.* **2015**, *50*, 377–404. [CrossRef]
24. Brownback, A.; Sadoff, S. Improving College Instruction through Incentives. *J. Political Econ.* **2020**, *128*, 2925–2972. [CrossRef]

25. Murayama, K.; Kitagami, S. Consolidation power of extrinsic rewards: Reward cues enhance long-term memory for irrelevant past events. *J. Exp. Psychol. Gen.* **2014**, *143*, 15–20. [CrossRef] [PubMed]
26. Murayama, K.; Matsumoto, M.; Izuma, K.; Matsumoto, K. Neural basis of the undermining effect of monetary reward on intrinsic motivation. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 20911–20916. [CrossRef]
27. Ding, X.-H.; He, Y.; Wu, J.; Cheng, C. Effects of positive incentive and negative incentive in knowledge transfer: Carrot and stick. *Chin. Manag. Stud.* **2016**, *10*, 593–614. [CrossRef]
28. Ferreira, M.; Cardoso, A.P.; Abrantes, J.L. Motivation and Relationship of the Student with the School as Factors Involved in the Perceived Learning. *Procedia-Soc. Behav. Sci.* **2011**, *29*, 1707–1714. [CrossRef]
29. Wilson, T.D. Models in information behaviour research. *J. Doc.* **1999**, *55*, 249–270. [CrossRef]
30. Kjaerulff, U.B.; Madsen, A.L. Bayesian networks and influence diagrams. *Springer Sci. Bus. Media* **2008**, *200*, 114.
31. Huang, Z.; Yang, L.; Jiang, W. Uncertainty measurement with belief entropy on the interference effect in the quantum-like Bayesian Networks. *Appl. Math. Comput.* **2019**, *347*, 417–428. [CrossRef]
32. Suter, P.; Kuipers, J.; Beerenwinkel, N. Discovering gene regulatory networks of multiple phenotypic groups using dynamic Bayesian networks. *Brief. Bioinform.* **2022**, *23*, bbac219. [CrossRef] [PubMed]
33. Anderson, R.D.; Mackoy, R.D.; Thompson, V.B.; Harrell, G. A Bayesian Network Estimation of the Service-Profit Chain for Transport Service Satisfaction. *Decis. Sci.* **2004**, *35*, 665–689. [CrossRef]
34. Anderson, R.D.; Vastag, G. Causal modeling alternatives in operations research: Overview and application. *Eur. J. Oper. Res.* **2004**, *156*, 92–109. [CrossRef]
35. Gupta, S.; Kim, H.W. Linking structural equation modeling to Bayesian networks: Decision support for customer retention in virtual communities. *Eur. J. Oper. Res.* **2008**, *190*, 818–833. [CrossRef]
36. Díez-Mesa, F.; de Oña, R.; de Oña, J. Bayesian networks and structural equation modelling to develop service quality models: Metro of Seville case study. *Transp. Res. Part A Policy Pract.* **2018**, *118*, 1–13. [CrossRef]
37. Kalisch, M.; Bühlmann, P. Estimating high-dimensional directed acyclic graphs with the PC-algorithm. *J. Mach. Learn. Res.* **2007**, *8*, 613–636.
38. Bianchi, P.; Elgui, K.; Portier, F. Conditional independence testing via weighted partial copulas. *J. Multivar. Anal.* **2023**, *193*, 105120. [CrossRef]
39. Ramirez-Hereza, P.; Ramos, D.; Toledano, D.T.; Gonzalez-Rodriguez, J.; Ariza-Velazquez, A.; Doncel, N. Score-based Bayesian network structure learning algorithms for modeling radioisotope levels in nuclear power plant reactors. *Chemom. Intell. Lab. Syst.* **2023**, *237*, 104811. [CrossRef]
40. Sun, B.; Zhou, Y.; Wang, J.; Zhang, W. A new PC-PSO algorithm for Bayesian network structure learning with structure priors. *Expert Syst. Appl.* **2021**, *184*, 115237. [CrossRef]
41. Jately, V.; Azzopardi, B.; Joshi, J.; Venkateswaran, V.B.; Sharma, A.; Arora, S. Experimental Analysis of hill-climbing MPPT algorithms under low irradiance levels. *Renew. Sustain. Energy Rev.* **2021**, *150*, 111467. [CrossRef]
42. Margaritis, D. *Learning Bayesian Network Model Structure from Data*; Carnegie-Mellon Univ Pittsburgh Pa School of Computer Science: Pittsburgh, PA, USA, 2003.
43. Tarka, P. An overview of structural equation modeling: Its beginnings, historical development, usefulness and controversies in the social sciences. *Qual. Quant.* **2018**, *52*, 313–354. [CrossRef] [PubMed]
44. Deng, L.; Yuan, K.-H. Multiple-Group Analysis for Structural Equation Modeling with Dependent Samples. *Struct. Equ. Model. A Multidiscip. J.* **2015**, *22*, 552–567. [CrossRef]
45. Bollen, K.A. *Structural Equations with Latent Variables*; John Wiley & Sons: Hoboken, NJ, USA, 1989; Volume 210.
46. Hu, L.T.; Bentler, P.M. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct. Equ. Model. A Multidiscip. J.* **1999**, *6*, 1–55. [CrossRef]
47. Bentler, P.M. Comparative fit indexes in structural models. *Psychol. Bull.* **1990**, *107*, 238. [CrossRef] [PubMed]
48. Bentler, P.M.; Bonett, D.G. Significance tests and goodness of fit in the analysis of covariance structures. *Psychol. Bull.* **1980**, *88*, 588. [CrossRef]
49. Zhang, L. Age matters for girls: School entry age and female graduate education. *Econ. Educ. Rev.* **2022**, *86*, 102204. [CrossRef]
50. Garaus, C.; Furtmüller, G.; Güttel, W.H. The Hidden Power of Small Rewards: The Effects of Insufficient External Rewards on Autonomous Motivation to Learn. *Acad. Manag. Learn. Educ.* **2015**, *15*, 45–59. [CrossRef]
51. Zhao, F.; Zhang, Z.-H.; Bi, L.; Wu, X.-S.; Wang, W.-J.; Li, Y.-F.; Sun, Y.-H. The association between life events and internet addiction among Chinese vocational school students: The mediating role of depression. *Comput. Hum. Behav.* **2017**, *70*, 30–38. [CrossRef]
52. Ong, M.; Smith, J.M.; Ko, L.T. Counterspaces for women of color in STEM higher education: Marginal and central spaces for persistence and success. *J. Res. Sci. Teach.* **2018**, *55*, 206–245. [CrossRef]

education sciences

MDPI

Article

Make Lectures Match How We Learn: The Nonlinear Teaching Approach to Economics

Peng Zhou

Cardiff Business School, Cardiff University, Cardiff CF10 3EU, UK; zhoup1@cardiff.ac.uk

Abstract: This paper proposes a nonlinear teaching approach, based on learning theories in cognitive psychology, with a special focus on large-cohort economics modules. The fundamental rationale is to match the features of teaching with the nature of learning. This approach was implemented in an undergraduate economics module, which received qualitative feedback and quantitative evaluation. Formal econometric models with both binary and continuous treatment effects were developed and estimated to quantify the effects of the proposed approach. Evidence shows that the nonlinear teaching approach significantly improves the effectiveness and efficiency of the learning-teaching process but does not promote student attendance.

Keywords: nonlinear teaching approach; higher education; experimental action research; treatment effects

JEL Classification: A22

Citation: Zhou, P. Make Lectures Match How We Learn: The Nonlinear Teaching Approach to Economics. *Educ. Sci.* **2024**, *14*, 509. https://doi.org/10.3390/educsci14050509

Academic Editors: Jian-Hong Ye, Yung-Wei Hao, Yu-Feng Wu and Savvas A. Chatzichristofis

Received: 31 March 2024
Revised: 23 April 2024
Accepted: 7 May 2024
Published: 9 May 2024

1. Introduction

The role of higher education in economic growth is well-recognised in the literature of human capital [1,2]. However, as the direct technology of human capital production, the teaching approach in higher education institutions seems to have evolved slowly, especially for quantitative subjects like economics [3,4]. The COVID-19 pandemic posed new challenges when recorded or online lectures were delivered to large cohorts with limited real-time engagement and feedback [5]. The recent development of generative AI (e.g., ChatGPT) has further transformed higher education and the role of lecturers [6]. To adapt to these disruptive challenges, blended teaching has become popular in universities [7]. Nevertheless, it is usually recognised that existing pedagogies do not always fit students' learning process [8]. Specifically, an important feature of learning process is nonlinearity—learners do not acquire new knowledge in a one-way fashion. To acquire a new piece of knowledge, a typical learner needs to go back and forth many times to eventually understand it and internalise it as part of their knowledge system. New knowledge can be confusing and is easily forgettable, especially when learners are still unclear about the general framework.

To illustrate how people learn, consider a visitor, Sue, who arrives at a new city, say, Cardiff. To know the city, she needs to walk around many times to accumulate information of the roads, districts, the places of interest, and so on. If Sue gets stuck in an unfamiliar corner and gets lost, the most natural choice for her is to try to get back to a place she is familiar with. After some repetition and iterations, a map is formed in Sue's mind. Once a general map is formed, Sue will be able to apply her knowledge of this city to guide her route. Similarly, learning a new theory has the same nonlinear feature. The first level of learning is to become familiar with the knowledge ("city"), corresponding to the process of drawing the map. The second level of learning is to apply the knowledge to practice ("route"). However, the two levels are usually not sequential but iterative. In both stages, learners may make mistakes and forget previous information—becoming "stuck in an unfamiliar corner". The reflection process will help learners to draw a better map and form better applications. An experienced learner may need less time and make less detours before she acquires the knowledge. Going back to our visitor's example, she may have already travelled to many similar cities before and know what a typical British city looks like. Then, she may get to know the city quite fast within one or two days thanks to her previous experience—"The learners learn how to learn as they learn." [9]

However, conventional teaching practices do not match the nonlinear nature of the learning process well. In some extreme cases, lecturers totally ignore how people learn. A typical lecture tends to be delivered like this: define the key concepts, outline the assumptions, derive the conclusions, and then apply to some popular contexts. As a well-known example in teaching undergraduate microeconomics, lecturers usually define the utility function, outline the maximisation problem, take derivatives, and then apply this information to some consumer problems. A national survey on teaching undergraduate economics was conducted in both 1996 and 2001 by Becker and Watts, who show that the main feature of economics teaching is still "Chalk and Talk" in the US. In contrast to the passive learning environment that characterises the teaching of economics, class discussion and other forms of active learning are now the dominant forms of instruction in other fields of higher education, rather than extensive lecturing [10].

There are two main reasons as to why economics lecturers are reluctant to use alternative teaching methods. On the one hand, economics by nature is a social science, heavily reliant on mathematics and logical derivations. This feature restricts the teaching methods in a passive fashion, similar to natural science disciplines. Active forms, such as discussion, tend to be avoided to achieve efficiency in teaching. However, efficiency does not imply effectiveness. The lecturer may well cover all the technical details of deriving some economic theory from assumptions to conclusions, but the students never know why they are doing these. An even worse outcome is, once you are confused during a lecture, you will never get back on track, because the lecturer is carrying on in a one-way direction. On the other hand, there are insufficient incentives in place to move instructors away from the status quo. For example, lecturers may try not to outperform colleagues and avoid being "too popular" with students to save time for their own research. In economic terms, there are both benefits and costs in teaching innovation, and the equilibrium depends on the optimal balance between the two. The second point can only be resolved by institutional changes in lecturer performance evaluation and is beyond the scope of this paper. Instead, the focus is on the "technological" aspect of improving the effectiveness and efficiency in teaching economics.

To see the drawbacks of linear teaching, let us go back to the visitor example. Assume Sue wants to visit Cardiff Castle from her hotel and she turns to a local resident, Joe, for help. If Joe uses linear teaching, then he may tell Sue to turn left in 100 yards onto North Road, turn right in 200 yards on to Colum Road, then enter the roundabout and take the third exit, etc. Unless Sue is a genius, she will soon get lost and frustrated, because the new road names are easily forgettable. It is not easy to follow all these instructions for those who do not even know where they are standing.

This ineffectiveness of teaching could be due to a lack of the "threshold concept", which is transformative and irreversible [11]. The transformative feature implies that before the learner grasps the threshold concept, she needs to accumulate information piece by piece and an understanding of the knowledge may be partial and erroneous. New information is always easy to forget and confuse. Once the learner has acquired the threshold concept, the general picture becomes clearer, and the learning curve will be steeper. The transformative feature entails a nonlinear teaching process to emphasise and reinforce the threshold concept for the learner. However, the irreversible feature of threshold concept brings difficulty in doing this. For those who have already internalised the threshold concept, they cannot recall how it felt when they did not. For the visitor example, Joe may think it is extremely easy to follow the instructions because he has travelled in that area for years. However, for Sue, it may be extremely difficult to follow because that area sounds like a black box to her. This makes communication difficult because lecturers do not even know where the students are confused.

This paper proposes an approach suitable for undergraduate, large-cohort economics teaching, though this approach can be easily applied to any social science subject at other levels. It is not conventional passive linear teaching (ineffective for economics), and it is not fancy active unstructured teaching (inefficient for economics) either. Rather, it cuts in the middle, which I refer to as the "nonlinear" teaching, because it appreciates the nonlinear nature of the learning process. The notion of nonlinear teaching offers an open system in which interconnected knowledge and sense-making are constantly re-built by dynamic interplays between students' existing knowledge and the new knowledge [12].

In fact, there are some practices in the literature sharing the same rationale of nonlinear teaching proposed here. For example, an early paper by Bartlett and King proposes teaching economics as a laboratory science, with the help of computers and software [13]. This approach emphasises learning-by-doing, i.e., the students are effectively "doing" economics like those in real economics profession. Hadsell proposes to promote students' engagement or active learning by confronting their own values and judgements against the economic theory [14]. This is also one way of sense making, because students only learn something when knowledge is accommodated into their existing system of belief. Other similar examples include Salemi [15] and Chow et al. [16], who promote active learning through group discussion or game play. All these attempts follow either a Behaviourist view or a Constructionist view, which form the starting point of this paper. However, the nonlinear teaching approach will go much further in associating the new theory and skills with the students' existing knowledge, so that the new knowledge can fit into their current knowledge framework. On the other hand, due to the disciplinary restrictions of economics, the proposed approach is not a radical revolution towards, say, totally unstructured open discussions in arts or philosophy. It takes into account the trade-off between effectiveness and efficiency in teaching and learning economics.

The contribution of this paper is to provide a formal theoretical foundation and a systematic teaching approach toolbox. Empirically, both qualitative and quantitative evaluations are conducted through an experimental action research, which has seldom been performed before in the literature. Following this introduction, Section 2 critically reviews the current literature and identifies gaps in large-cohort teaching in higher education. Section 3 discusses the conceptual framework and the theoretical rationale behind the nonlinear teaching approach. Section 4 introduces the nonlinear teaching intervention in practice. Section 5 provides some reflections on the qualitative feedback and a formal quantitative evaluation of the nonlinear teaching practice using econometric models. Section 6 concludes.

2. Literature Review

Teaching large cohorts in higher education, particularly in quantitative fields like economics, presents distinct challenges and opportunities for pedagogical innovation. Traditional lecture methods often do not sufficiently engage students or accommodate diverse learning styles, which can impede the understanding and retention of complex economic concepts [17]. This review examines various teaching approaches designed for large classes and discusses the various indicators measuring the effectiveness and efficiency of teaching approaches.

2.1. Large-Cohort Teaching

Conventionally, large-cohort teaching has relied heavily on lecture-based methods, characterised by one-way communication from the instructor to the students [18]. This model has been criticised for promoting passive learning, where limited interaction can lead to decreased student engagement and poor retention [19]. In response, educational strategies have evolved to incorporate more interactive and student-centred approaches [20]. For example, the flipped classroom model reverses traditional expectations by requiring students to prepare before class, thus enabling active learning during lecture sessions through problem-solving and discussion [21]. Technology-enhanced learning tools, such as clickers or online forums, have also been integrated to facilitate real-time feedback and increase student participation [22].

The effectiveness of large-cohort teaching is significantly enhanced when aligned with established learning theories [23]. Constructivist learning theory, which posits that learners construct their understanding and knowledge of the world through experiences and reflection on those experiences, supports environments that actively involve students in their learning [24]. Applying constructivist principles, instructors can facilitate learning by encouraging students to question, explore, and apply ideas in real-world contexts [25]. Experiential learning theory further supports this approach by emphasizing the role of experience in the learning process [26]. In large economics classes, these theories have been operationalised through case-based teaching where students analyse real economic scenarios, and through simulation games that mimic economic decision-making processes [27]. Research indicates that such practices not only improve conceptual understanding but also enhance students' ability to apply economic theories pragmatically [28].

Studies evaluating the impact of innovative teaching methods in large cohorts provide mixed results. Freeman et al. conducted a meta-analysis showing that active learning significantly increased student performance in science, engineering, and mathematics [19]. Similar studies in economics suggest that, while active learning techniques like the flipped classroom can improve examination scores, they require substantial adaptation by instructors and commitment from students [29]. Additionally, the use of technology in large classes has been shown to facilitate interaction and engagement but also presents challenges related to distraction and the digital divide [30].

After all, these innovative teaching approaches are only complementary to, not substitutes for, lectures [31]. Few attempts have been made to improve the effectiveness of learning and teaching within the traditional form of lectures. Essentially, the literature on large-cohort teaching mainly gets around the question by avoiding lectures, rather than resolving the problem per se. This paper, in contrast, aims to fill the gap in the literature directly. Hence, this paper aims to answer two research questions:

RQ1. How to make economics lectures match how we learn?
RQ2. How to evaluate the proposed nonlinear teaching approach?

2.2. Indicators for Evaluating Teaching Approaches

The effectiveness and efficiency of teaching approaches are paramount to achieving educational objectives. Therefore, indicators measuring effectiveness and efficiency are informative tools to evaluate different teaching approaches. There are three prevailing indicators in empirical literature.

First, student performance, often measured through marks, is a traditional and powerful indicator of the effectiveness of teaching methods. In economics education, where quantitative and analytical skills are emphasised, marks not only reflect the acquisition of subject-specific knowledge, but also critical thinking and problem-solving abilities. Studies have demonstrated a correlation between innovative teaching methods, such as case-based learning and simulation, and improved student performance in economics courses [32]. These findings suggest that methods that actively engage students tend to enhance their understanding and retention of complex economic theories [33].

Second, attendance rate is also frequently used as an indirect measure of student engagement and the attractiveness of the teaching approach. High attendance rates are often associated with more engaging lectures and a positive classroom environment, which are critical in subjects as challenging as economics. Research by Gupta and Pandey indicates that interactive lectures, where students participate in discussions and problem-solving activities, significantly boost attendance compared to traditional lecture-based sessions [34]. This correlation underscores the importance of interactive elements in maintaining student interest and engagement.

Third, student satisfaction surveys are a vital tool for assessing the quality of teaching and the overall student experience. These surveys typically evaluate various aspects of the course and teaching methods, including the clarity of instruction, the relevance of the material, and the instructor's ability to inspire interest in the subject. In the field of economics, where theoretical concepts can seem detached from practical realities, teaching methods that effectively bridge this gap, such as real-world applications and experiential learning, tend to score higher on satisfaction metrics. A study by Ang et al. highlights that students rate courses higher when instructors successfully linked economic theories to current events and real-life economics issues [35].

Beyond marks, attendance, and student satisfaction, the literature on educational assessment recognizes several other indicators that can be used to measure the effectiveness and efficiency of teaching approaches. For example, graduation and retention rates are crucial indicators, particularly in higher education. Effective teaching methods should not only engage students, but also support them in successfully completing their courses and programs. Higher retention and graduation rates are often seen as indicators of successful teaching strategies that foster both academic and personal development among students [36]. In addition, as the ultimate aim of many educational programs, especially in higher education, is to prepare students for successful careers, the rate at which students gain employment in their field of study or advance in their careers can be an indicator of the effectiveness of teaching. This metric evaluates how well educational content and teaching methods prepare students for the professional world [37]. Moreover, feedback from alumni about the long-term impact of their education can provide valuable insights into the effectiveness of teaching methods. Alumni surveys might ask about the relevance of the skills they learned, their preparedness for professional challenges, and their overall educational satisfaction. This long-term perspective helps institutions to understand the enduring impact of their teaching methods [38]. Nevertheless, these indicators are more appropriate for the programme-level evaluation rather than lecture-level evaluation. As a result, we will adopt the three traditional indicators, i.e., marks, attendance, and satisfaction, in our empirical section. The next section will develop the proposed nonlinear teaching approach based on learning theories.

3. Conceptual Framework

The theoretical rationale of the nonlinear teaching approach is learning theories. The basic idea of this approach is to match the teaching and learning process and to promote the effectiveness and efficiency of disseminating and acquiring knowledge. There are several schools of thought on learning process. Learning theories provide a variety of conceptual frameworks to describe how information is absorbed, processed, and retained during learning. Cognitive, emotional, and environmental influences, as well as prior knowledge, all play a role in how knowledge and skills are acquired. There are two conventional learning theories shedding light on the proposed nonlinear teaching approach: Behaviourism and Constructionism.

3.1. Behaviourism

Behaviourism was coined by John Watson (1878–1959), who argued that learning is an aspect of conditioning, and advocated for a system of rewards and targets in education. In a simplified interpretation, Behaviourists model the learning process as stimulus—response–reinforcement [39]. If the lecturer can correctly provide some reward system to motivate the students, the effectiveness of the learning process may be promoted. Responses that result in favourable outcomes tend to be repeated and become established behaviour, while responses that lead to negative or neutral outcomes tend not to be repeated.

This learning-by-doing nature of the learning process implies that teaching process should provide appropriate stimulus from time to time. However, in a conventional "linear" teaching approach, as illustrated in Figure 1, the lecture contents usually start from simple things and become more and more difficult. The degree of complicatedness monotonically increases as the lecturer builds new knowledge on the previous elements. This linear fashion tends to make students frustrated (negative stimulus) in the middle of the module.

Once lost, always lost. It argues for a nonlinear design of the module contents in terms of difficulty, so that the students are rewarded by being able to understand (positive stimulus).

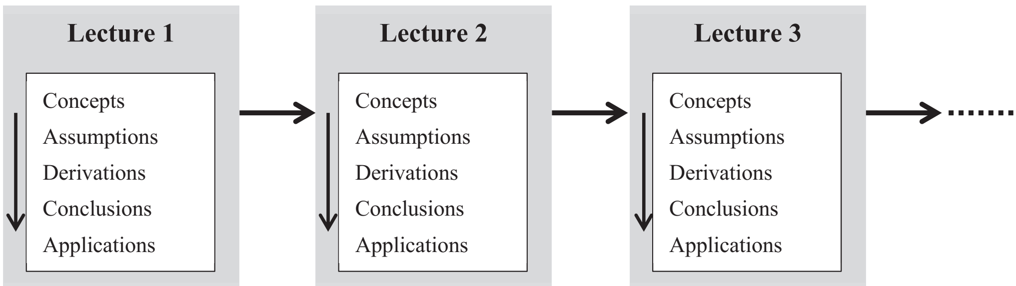


Figure 1. Linear Teaching Conceptual Framework.

There seems to be a conflict between the logical structure and the nonlinear design of the lecture contents. Logically speaking, the module should be designed accumulatively, i.e., later lectures should be built on the previous lectures. It inevitably leads to a simple-to-difficult linear fashion. However, this logical structure is an outcome of the learning process, rather than the learning process per se. No theory has been born as what it eventually looks like. There are detours and errors during the evolution of the theory, and the learning process is the same. Therefore, the teaching procedure does not have to be "logical" and precise in every stage. To provide a learning experience with positive stimulus, the lecturer can use some naïve interpretations (easy to understand but imprecise or even erroneous) in the first stage and correct them to a more precise definition later on (as illustrated in Figure 2). Moreover, nonlinear teaching is not from specific to general (as shown in linear teaching), but from general to specific in a holistic fashion. Of course, the lecturer must make the students aware that the initial interpretation is not precise and to be corrected and detailed later on. In this way, a module may have several lectures, in each of which we start with a simple and general understanding while continually correcting the initial understanding to a more precise level, so that the learners receive positive stimulus constantly while acquiring the knowledge. The idea behind this is that there is no shortcut in learning. Making mistakes is both natural and necessary for learning, and some appropriate detour may actually be more effective than linear progress.

To provide intuition, we revisit our example in the introduction again. To teach Sue to become familiar with the city, the local guide Joe may start like this. The city centre is very close to the Millennium stadium, which is very tall and easy to see from anywhere (initial stage). Follow the direction which leads you closer to that landmark and adjust when you are approaching the Millennium stadium (correcting). It is very likely that Sue will come across the city centre during the journey towards the stadium. Once this task (going to city centre) is accomplished, Joe can then teach Sue the next "lecture", say, going to national museum. A similar procedure can then be used.

However, Behaviourism does not account for free will and the internal prior knowledge of the learners. Therefore, the nonlinear teaching approach also draws some inspirations from Constructionism.

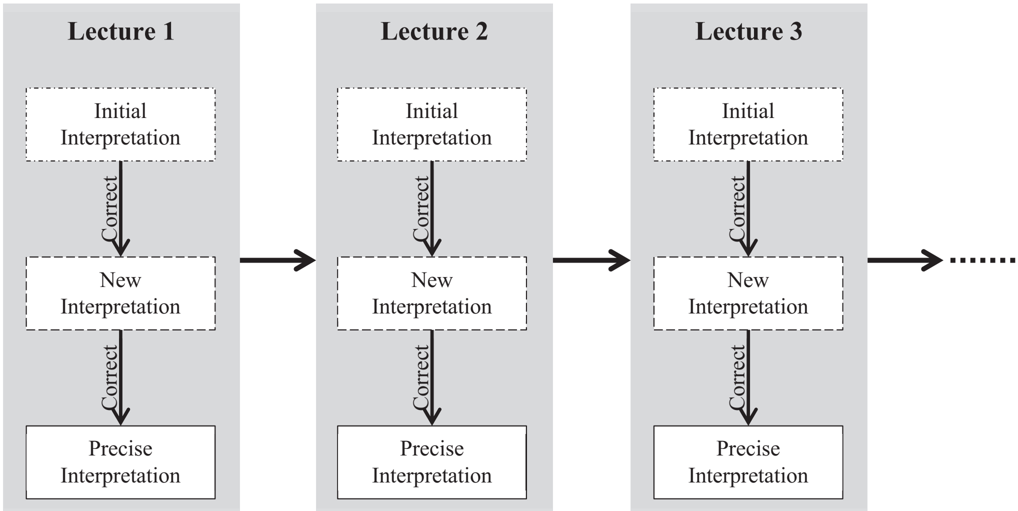


Figure 2. Nonlinear Teaching Conceptual Framework I.

3.2. Constructionism

Constructionism treats learning as a process of sense making, so the role of the instructor is as the facilitator. The two pioneer Constructionists, Piaget and Vygotsky, both appreciated the essence of internalising knowledge, rather than accepting the information as presented through rote-memory [40,41]. Everyone comes to the lecture with some existing knowledge, and the learning process is to accommodate the new information into the old system and framework. This implies that the lecturer needs to take into account the acceptability and conflicts between the contents of lecture with learners' prior knowledge.

This also implies that learning is not a one-way process. Rather, learners go back and forth between the new information and the existing knowledge. To match this nonlinear feature of learning process, lecturers need to bridge the gap between the learners' existing knowledge and the new knowledge. This connection should be built from time to time, not just once in the beginning of the lecture, because this feature of the learning process is continuous.

Let us still use the visitor's example, when Joe teaches Sue how to go to the national museum. Joe can always base his teaching on Sue's existing knowledge about the city. For example, now that Sue already knows how to go to the city centre, Joe can just tell Sue that the museum is just to the north of the city centre. The existing knowledge provides a good reference point to compare the new knowledge with.

One difficulty of associating with prior knowledge in teaching is that different students may have different backgrounds, and so, different prior knowledge. Sometimes, they do not even have prior knowledge, especially for introductory level modules. In this case, we can still use common sense or even metaphors to establish the bridge. It proves quite helpful for students to understand the abstract theories if lecturers have a good sense of humour. One explanation for this is because humour can link the complicated and dry theory to some vivid everyday experience which shares similar logic. For example, in macroeconomics, lecturers usually describe the relationship between the government and the public as a couple. The government may promise some policy during the election but follow a different policy afterwards (so-called "time inconsistency"). This can be likened to how a boy promises to love the girl during courting, but after establishing the relationship, the boy may change his behaviour. This sort of metaphor may greatly assist the students to understand the essence of theories, though the description is not precise. Another by-product of using everyday examples to provide sense making is that students enjoy the learning process more and attendance could be improved.

Therefore, the sense making does not change the conceptual framework of the nonlinear teaching method developed above, but it provides an effective way of raising the initial interpretation of new knowledge. Usually, it can be a bird's eye-view of "where we are" from the whole picture. It is suggested to place the new knowledge into a bigger context, with which students are already familiar. Comparison and contrast with familiar contents can be used to help students accommodate the new knowledge in an existing framework. This can be supplemented by everyday examples or common sense to help students to grasp the essence of the theory. After this, a more systematic and precise exposition of the theory should be laid out and summarised.

3.3. Nonlinear Teaching Approach

In the light of Behaviourism and Constructionism, I developed the general framework of the nonlinear teaching approach, as well as some techniques of implementing it. Other learning theories, such as social learning, may also contribute to enhancing the effectiveness and efficiency of teaching and learning. For example, assigning some roles for the students to solve some realistic problems. Group discussion can also be used to promote the exchange of knowledge and understanding in a multi-directional fashion, rather than linear one-directional flow from lecturer to students.

The nonlinear teaching approach can be formulated in a more precise version as in Figure 3, which is an upgraded and revised framework from Figure 2. The key difference is the review and reflection upon the system of knowledge along with the learning and teaching process, instead of just once at the beginning of each lecture and then a linear flow within each lecture. The purpose of revisiting "where are we" is to make sure the learners keep a close track of the teaching plan about what have been taught and what is to be taught. It can avoid the difficulty in finding their way once lost.

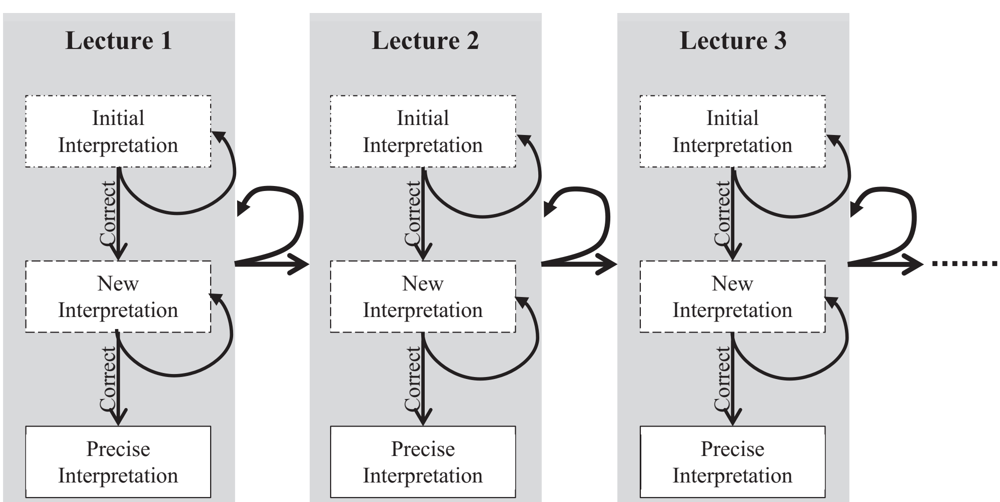


Figure 3. Nonlinear Teaching Conceptual Framework II.

One feasible and simple technique to realise this additional nonlinear feature in teaching is to show the outline after each section of each lecture. It could be even better if lecturers add a navigation panel in their slides. It is also advisable to revisit some important and/or difficult points, even if you think it is a bit repetitive. It is not repetitive to the learners!

Let us turn to our visitor's example once again. In addition to telling the route to Sue, Joe also gives Sue a GPS navigator, which can show her where she is in the city. That would greatly enhance the learning process. Even if Sue gets lost due to some confusing lanes, she will be able to get back to the right direction as long as she looks at the navigator. The outline in the lecture is like a static map, which helps, but the revision of the outline is like a dynamic GPS navigator, which helps a lot more.

To summarise, the nonlinear teaching approach is mainly derived from Behaviourism and Constructionism. The rationale is to match the nature of the learning process in teaching. We also discussed some techniques to implement nonlinear teaching including, but not restricted to:

- teaching from a rough to precise perspective;
- teaching from a general to specific perspective;
- sense-making, based on common knowledge and metaphors;
- role play and group discussion to promote active and interactive learning;
- a review of "where are we" at the beginning of each section and each lecture.

Note that nonlinear teaching does not mean simple repetition, but a spiral ascending process. The same contents may be recalled several times, but it should deepen the learners' understanding each time.

In fact, the readers may have already realised that this paper is written in a nonlinear teaching approach style. I started with an everyday life example (visitor) and have revisited this example several times. The illustration of the conceptual framework started with a simplified rough idea in Figure 2 and then became more precise by Figure 3. Imagine that if Figure 3 is shown directly upfront, some readers may feel confused due to the complicated lines and curves. A small step further with many steps may work much better than one big stride. If you have understood this paper well up to now, then the nonlinear teaching approach works on you.

4. Teaching Intervention in Practice

I implemented the nonlinear teaching approach in Intermediate Microeconomics, a large-cohort undergraduate core module every year in a British university. The following techniques were used.

First, in each lecture slide, I presented a navigation tree, showing where we are. I explicitly revisited the outline at the beginning of each section. Moreover, in the beginning of each lecture, I also revisited the topics covered and posited the current lecture in the big picture drawn in the first lecture. A simple revision was provided on each part when it was finished, but from a more general and higher perspective.

Second, everyday examples were elaborately designed to assist the students in accommodating the new theory into their existing knowledge system. For example, when I described "production function" in producer theory, a metaphor is used. I told the students to treat their exam marks as the "output" from the production function, and the students' effort as one input, while attendance as another input. More effort and a higher attendance tended to bring higher marks. A smart student with low inputs may still achieve a lower mark than another diligent student with higher inputs. The intelligence is the curvature of the production function, while the diligence is the inputs of the production function. Later on, this example can be used again to illustrate the concept of a "diminishing marginal product". For a given level of effort, a student's mark rises less and less if only one type of input rises.

Third, as mathematics is a main feature in Intermediate Microeconomics, such as differentiation and equation solving, to avoid confusing the students in the middle of the lecture, I always presented the intuition first, then the mathematical details, as well as a graphical illustration. More than one approach was presented to students, so that those with a weaker mathematical background could still follow the general argument without getting lost in the rest of the lecture.

Finally, group discussion was frequently used in seminars. Interactions among students were greatly encouraged. When one's role shifts from a learner to a discussant, students tend to be more active and understand the knowledge deeper in communication, because there is more than one angle of looking at the problems (Zhou, 2019). This method replaces uni-directional linear information flows from the lecturer to students (Figure 4A). Multi-directional nonlinear information flows among the lecturer, students, and other sources like the Internet and AI-based tools (Figure 4B).

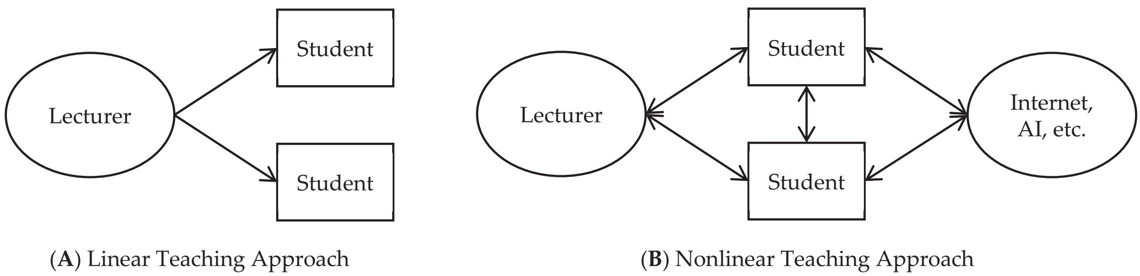


Figure 4. Information flow of linear and nonlinear teaching approaches.

The next section will evaluate the effectiveness and efficiency of the nonlinear teaching approach using strict econometric models.

5. The Evaluation of the Approach

As a lecturer, I have access to the students' information, such as gender, race, age, programme, exam marks, and attendance for both current students (the "treated" group) and previous students (the "control" group). The evaluation of the nonlinear teaching intervention is based on two types of feedback. Conventional feedback was collected through questionnaires at the end of the semester before an exam. The response rate was 85.7%. However, qualitative evaluation based on a questionnaire is usually criticised as being biased because the students may fear that the lecturer can figure out who wrote what. Also, the answers to the questionnaire may only reflect what they think is true, not what is actually true. To make a consistent comparison between the "treated" students and the "control" students, I collected the second type of "feedback" by conducting a mock exam with the current students using the previous year's exam paper. This is like a natural experiment with the exogenous assignment of treatment, and it is guaranteed to compare like with like. One may argue that the current students may not be as well prepared when they take the mock exam as they are taking their actual exam. It could be handled by using a weighted average between their actual and mock exam marks. However, this remedy also has its drawbacks, for example what weight should be used. Since the mock exam is conducted about three weeks before the actual exam, it is not implausible to assume the students have a consistent performance. Therefore, it is complementary and superior to use the second type of feedback, based on their actual performance in their exam (or mock exam) marks. It is more objective and can provide a quantitative measure for effectiveness and efficiency. Therefore, I will briefly discuss the qualitative evaluation based on the first type of feedback and detail the quantitative evaluation based on the second type of feedback.

5.1. Qualitative Evaluation

The questionnaire asks the students two free-text questions:

(i) What are you finding most useful about the module?
(ii) What would improve the module?

For the first question, almost all students (94.4%) noted that the lectures were "well-structured" or "easy to follow", which implies that the navigation of the big picture improved the students' learning experience. About a half of them (47.2%) mentioned that the lectures were closely relevant to "everyday life" and a "working environment", indicating that sense-making based on common knowledge worked well. Moreover, some students appreciated the multiple ways of presenting the same theory, including graphical, mathematical, and verbal. This nonlinear feature of the teaching approach also helps students from different backgrounds to understand the theory better and deeper.

Note that students were not aware that the lecturer carried out such a nonlinear teaching intervention before they filled in the questionnaire. This is to avoid biasing their opinions. Surprisingly, they seemed to identify the techniques of nonlinear teaching, and this is supportive evidence for the effectiveness of the teaching intervention.

For the second question, a main complaint is about the mathematics used in the module. Some business students do not have A-level maths, so they felt a bit challenged following the derivation of some mathematical models. This will be addressed by organising different seminar groups in the future.

5.2. Quantitative Evaluation

In addition to the qualitative evaluation, this paper also designs a formal econometric model to quantify the effectiveness and efficiency of the nonlinear teaching approach based on the second type of feedback. The mock exam marks of the current students are compared to the actual exam marks of the previous students, after controlling for other individual characteristics. There are two advantages of the teaching intervention. On the one hand, the outcome (the marks) is independent of the assignment of "treatment" by design, so it is a "natural experiment" (i.e., the data is experimental rather than observational). Thus, the treatment effect can be directly estimated without having to worry about endogeneity or a selection bias problem. On the other hand, the students have different attendance rates to lectures and seminars, which makes the "treatment" a continuous variable. Those who do not attend some of the lectures or seminars must review the contents themselves, so the teaching intervention has less effect on them. It is assumed that the students are occasionally absent because of random events, such as sickness or emergency, rather than a systematic lazy personality (In fact, the attendance is strictly monitored by the school and the students are required to provide justifiable reasons for absence. However, some students who consistently have low attendance (lower than 20%) are dropped from the analysis, because they may present a systematic difference in motivation). Therefore, the assignment of treatment is still exogenous, but this enables us to explore the treatment effect in more details. Nevertheless, the use of mock exams as a metric for learning outcomes assumes that students treat these mocks with the same seriousness as actual exams. If this is the case, then the students sitting in the mock exam should perform less well than if they are sitting in actual exams. The estimated treatment effect is a lower bound, which reinforces our argument, rather than weakens it.

If we treat the students' marks as the "output" of a "production function" of human capital, with students' attendance (A_i) as the "input". The "total factor productivity" of the production function can be affected by both the students' individual characteristics, including gender, race, native/overseas student, with/without A-level maths and business/economics students, as well as the lecturer's teaching approach (the teaching intervention dummy, or the "treatment", T_i). The implied econometric model can be written as:

$$marks_i = \alpha + \mathbf{ind}_i \cdot \boldsymbol{\beta} + \gamma \times A_i + \delta \times T_i + \phi \times A_i \times T_i + \varepsilon_i$$

The intercept α is the average level of productivity of an average student, and the vector $\boldsymbol{\beta}$ captures the effects of individual characteristics ($\mathbf{ind}_i$) on productivity. The coefficient γ can be interpreted as the return to attendance, i.e., how many marks an average student could improve by one more attendance. The coefficient δ is the "treatment effect" of the teaching intervention, because only the "treated" students ($T_i = 1$) have this shifting term in their productivity (intercept effect). Finally, there is also a cross-product term (the attendance rate A_i multiplied by the intervention dummy), and the coefficient ϕ is the effect of the treatment on the return to attendance (slope effect). If a student is exposed to the nonlinear teaching approach, then their return to attendance is equal to $\gamma + \phi$, rather than γ. The coefficient δ can be used to evaluate the effectiveness, while ϕ can be used to evaluate the efficiency, because the different degrees of exposure to the teaching intervention also improve learning.

In total, there are 342 students registered in the treated group, including both economics students (78%) and business students (12%). Three students have consistently low attendance and are dropped in the analysis. The control group has 316 students with a similar structure, but all students are included. The distribution of marks and attendance of the two groups are contrasted in Figure 5. It is shown that the distribution of marks of the treated group is more concentrated and more towards higher marks, and so is the distribution of attendance.

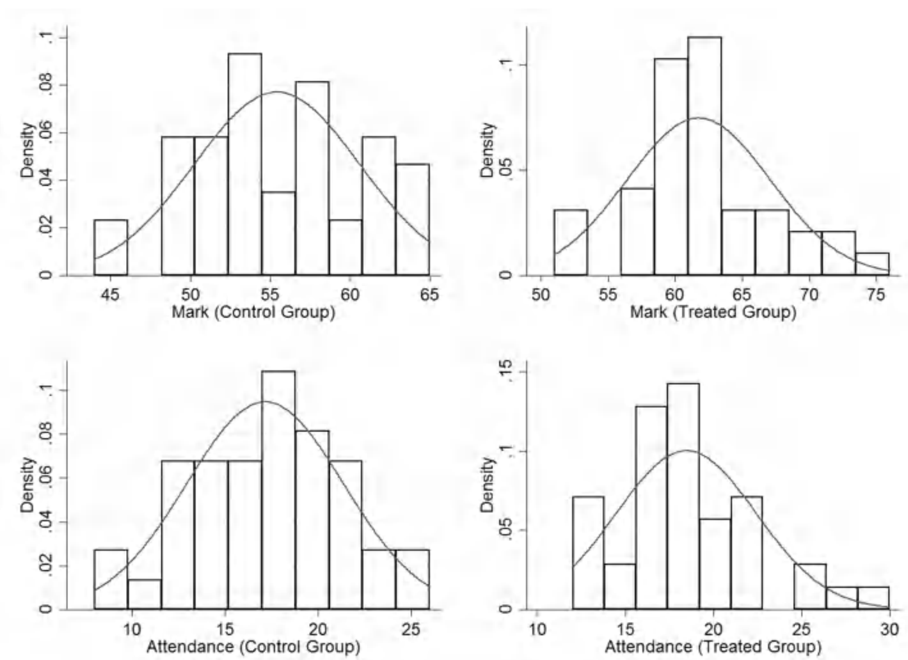


Figure 5. Distribution of Marks and Attendance.

The regression results shown in Table 1 are a stricter way to verify the observed changes in distributions of marks and attendance. The first two columns of the table focus on evaluating the effectiveness and efficiency of the nonlinear teaching approach using two alternative model specifications (both are estimated using OLS with robust standard errors against possible heteroskedasticity), and the last column attempts to find out if the new teaching intervention improves the attendance (Poisson regression is used because the dependent variable is a count variable). To summarise, the three hypotheses to be tested are:

Hypothesis (H1). *The nonlinear teaching approach increases the effectiveness of learning and teaching, or a positive average mark, i.e.,* $\delta > 0$.

Hypothesis (H2). *The nonlinear teaching approach increases the efficiency of learning and teaching, or a higher return to attendance, i.e.,* $\phi > 0$.

Hypothesis (H3). *The attendance rate is improved because of H1 and H2.*

Both model specifications imply a significant and positive treatment effect of the nonlinear teaching approach, so H1 is verified. If we do not include the cross-product term $A_i \times T_i$ as in the first model specification, then the treatment effect (the coefficient of T_i) is $\hat{\delta} = 7.428$. In other words, the students' performance (marks) is improved by more than seven marks on average. However, if the cross-product term is included, as in the second specification, the treatment effect reduces to $\hat{\delta} = 4.171$, but the return to attendance increases from 0.469 to $\hat{\gamma} + \hat{\phi} = 0.384 + 0.182 = 0.566$. This finding verifies H2. Based on the two regressions, we can actually decompose the overall treatment effect into two components: (i) the improvement purely due to a more effective teaching practice, accounting for $\frac{4.171}{7.428} = 56.15\%$ of the treatment effect; and (ii) the improvement due to the indirect effect on return to attendance, i.e., efficiency improvement, accounting for the rest, 43.85%.

Table 1. Estimation Results.

Dep. Var.	Mark	Mark	Attendance (A)
male	−0.511	−0.59	−0.04
	(0.462)	(0.456)	(0.06)
white	0.471	0.464	0.037
	(0.548)	(0.512)	(0.067)
native	1.062 **	1.021 **	0.005
	(0.446)	(0.468)	(0.066)
A-level	9.147 ***	9.138 ***	0.022
	(0.431)	(0.416)	(0.061)
business	−4.739 ***	−4.775 ***	−0.011
	(0.528)	(0.518)	(0.07)
A_i	0.469 ***	0.384 ***	
	(0.048)	(0.053)	
T_i	7.428 ***	4.171 ***	0.078
	(0.407)	(1.545)	(0.057)
$A_i \times T_i$		0.182 **	
		(0.088)	
_cons	44.132 ***	45.684 ***	2.828 ***
	(0.915)	(1.028)	(0.099)
Method	Robust OLS	Robust OLS	Poisson
R-sq	0.926	0.93	
adj. R-sq	0.919	0.922	
pseudo R-sq			0.007
AIC	323.482	321.519	462.419
BIC	342.538	342.957	479.093

Notes: Standard errors in the parentheses, ** 5%, *** 1%.

In particular, the efficiency improvement can be quantified by the coefficient of the cross-product term relative to that without treatment: $\frac{0.182}{0.384} = 47.40\%$. This significantly higher return to attendance should, in principle, encourage a higher attendance, because their marks can be improved more efficiently by attending the lectures (H3). This hypothesis is tested by a Poisson regression (the last column), but it turns out that the teaching intervention does not significantly increase the students' attendance. This is perhaps because the students are not aware of the implementation of the teaching intervention while it was implemented.

Other interesting findings from the regressions include:

- Gender and race do not significantly contribute to different marks.
- Native students tend to perform better than international students, maybe because of a language advantage.
- Students with A-level maths are expected to perform better than those without.

- Students enrolled in business programmes tend to obtain lower marks than those enrolled in economics programmes.

6. Conclusions

This paper proposes a nonlinear teaching approach in the context of teaching economics, though it is ready to be generalised to other disciplines. The rationale is derived from two learning theories, Behaviourism and Constructionism. By fitting the nature of the learning process, the nonlinear teaching approach is expected to improve the effectiveness and efficiency of teaching. A main feature of nonlinear teaching approach is to review and associate the new knowledge with the existing framework, either via sense making or via progressive repetition. It lies between the conventional passive teaching and the unstructured interactive teaching. Therefore, nonlinear teaching is more appropriate to natural science and quantitative social science like economics.

Some specific techniques are discussed and applied to a specific scenario in an undergraduate teaching practice. The teaching intervention received positive feedback in both qualitative and quantitative evaluation. Formal regression analysis verifies the effectiveness and efficiency of the nonlinear teaching approach based on the natural experiment (H1 and H2). It is estimated that 56.15% of the overall improvement on student performance can be attributed to the more effective teaching approach (an intercept effect), and the other half is derived from a higher return to attendance (a slope effect). If the students were aware of this, they would have improved their attendance, but this hypothesis (H3) is not supported in the Poisson regression.

One possible shortcoming of nonlinear teaching approach, as reflected by the student feedback, is that too much repetition may make the students bored and less receptive. This might explain why attendance was not improved significantly, because the negative effect cancels out the positive effect on the attractiveness of the lectures. Using the visitor's example again, if Sue goes through the same roads in Cardiff hundreds of times, she may lose her joy of travelling—which turns to boring commuting. If there is less passion, then there is less joy of travelling. Therefore, lecturers who adopted nonlinear teaching may have to think of a way of promoting the students' motivation and interest while benefitting from the improvement in effectiveness and efficiency.

It is also worth noting that there is no one-size-fits-all solution in teaching. Different types of modules require different types of teaching approaches. For example, for introductory modules like introduction to economics, teaching approaches often emphasize clarity, simplicity, and motivation. For advanced modules like microeconomic theory and macroeconomic theory, teaching approaches may need to be more structured. The proposed nonlinear teaching approach is more appropriate for the latter. The approach has been supported in a core module of undergraduate economics programmes, but its principles can be applicable to other subjects with similar features—abstract concepts and quantitative contents. In general, with the fast development of AI technology, the role of lecturers must evolve from a knowledge dispenser to a paradigm navigator.

Funding: The APC was funded by Open Access Support Team, Cardiff University.

Institutional Review Board Statement: Ethical review and approval were waived for this study because student feedback used in this paper was part of regular teaching and learning activities.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data and codes are available on request.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Lucas, R. On the Mechanics of Economic Development. *J. Monet. Econ.* **1988**, *22*, 3–42. [CrossRef]
2. Foreman-Peck, J.; Zhou, P. Late Marriage as a Contributor to the Industrial Revolution in England. *Econ. Hist. Rev.* **2018**, *71*, 1073–1099. [CrossRef]
3. Becker, W.E.; Watts, M. Chalk and Talk: A National Survey on Teaching Undergraduate Economics. *Am. Econ. Rev.* **1996**, *86*, 448–453.
4. Becker, W.E.; Watts, M. Teaching Economics at the Start of the 21st Century: Still Chalk-and-Talk. *Am. Econ. Rev.* **2001**, *91*, 446–451. [CrossRef]
5. Henne, A.; Möhrke, P.; Huwer, J.; Thoms, L.-J. Learning Science at University in Times of COVID-19 Crises from the Perspective of Lecturers—An Interview Study. *Educ. Sci.* **2023**, *13*, 319. [CrossRef]
6. Bahroun, Z.; Anane, C.; Ahmed, V.; Zacca, A. Transforming Education: A Comprehensive Review of Generative Artificial Intelligence in Educational Settings through Bibliometric and Content Analysis. *Sustainability* **2023**, *15*, 12983. [CrossRef]
7. Zheng, L.; Lee, K.C. Examining the Effects of "Small Private Online Course and Flipped-Classroom"-Based Blended Teaching Strategy on First-Year English-Major Students' Achievements. *Sustainability* **2023**, *15*, 15349. [CrossRef]
8. Rodríguez, M.F.; Nussbaum, M.; Pertuzé, J.; Avila, C.; Caceres, J.; Valenzuela, T.; Ceppi, B. Using metacognition to promote active learning in large business management classes. *Innov. Educ. Teach. Int.* **2021**, *59*, 410–420. [CrossRef]
9. Biehler, R.F.; Snowman, J. *Psychology Applied to Teaching*, 13th ed.; CENGAGE Learning Custom Publishing: Boston, MA, USA, 2011.
10. McCullough, K.; Munro, N. Finance students' experiences of lecture-based active learning tasks. *Innov. Educ. Teach. Int.* **2018**, *55*, 65–73. [CrossRef]
11. Cousin, G. An Introduction to Threshold Concepts. *Planet* **2006**, *17*, 4–5. [CrossRef]
12. Doll, W.A. Reflections on Teaching: Developing the Nonlinear. *Teach. Educ.* **1999**, *10*, 39–54. [CrossRef]
13. Bartlett, R.L.; King, P.G. Teaching Economics as a Laboratory Science. *J. Econ. Educ.* **1990**, *21*, 181–193. [CrossRef]
14. Hadsell, L. Exploring Values in the Classroom: Three Exercises for Introductory Economics. *Am. Econ.* **2005**, *49*, 51–59. [CrossRef]
15. Salemi, M.K. An Illustrated Case for Active Learning. *South. Econ. J.* **2002**, *68*, 721–731. [CrossRef]
16. Chow, J.Y.; Davids, K.; Button, C.; Shuttleworth, R.; Renshaw, I.; Araújo, D. Nonlinear pedagogy: A constraints-led framework for understanding emergence of game play and movement skills. *Nonlinear Dyn. Psychol. Life Sci.* **2006**, *10*, 71–103.
17. Bligh, D.A. *What's the Use of Lectures?* First US Edition of the Classic Work on Lecturing; John Wiley & Sons: Hoboken, NJ, USA, 2000.
18. Su, C.-Y.; Chen, C.-H. Exploring and comparing pedagogical beliefs of university instructors in relation to their behavioural patterns regarding learning management system use. *Behav. Inf. Technol.* **2024**, 1–14. [CrossRef]
19. Freeman, S.; Eddy, S.L.; McDonough, M.; Smith, M.K.; Okoroafor, N.; Jordt, H.; Wenderoth, M.P. Active learning increases student performance in science, engineering, and mathematics. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 8410–8415. [CrossRef] [PubMed]
20. Coleman, T.E.; Money, A.G. Student-centred digital game–based learning: A conceptual framework and survey of the state of the art. *High. Educ.* **2020**, *79*, 415–457. [CrossRef]
21. Lage, M.J.; Platt, G.J.; Treglia, M. Inverting the Classroom: A Gateway to Creating an Inclusive Learning Environment. *J. Econ. Educ.* **2000**, *31*, 30–43. [CrossRef]
22. Caldwell, J.E. Clickers in the Large Classroom: Current Research and Best-Practice Tips. *CBE—Life Sci. Educ.* **2007**, *6*, 9–20. [CrossRef]
23. Roßnagel, S.C.; Fitzallen, N.; Lo Baido, K. Constructive alignment and the learning experience: Relationships with student motivation and perceived learning demands. *High. Educ. Res. Dev.* **2021**, *40*, 838–851. [CrossRef]
24. Piaget, J. *The Construction of Reality in the Child*; Basic Books: New York, NY, USA, 1954.
25. Zajda, J. (Ed.) Constructivist Learning Theory and Creating Effective Learning Environments. In *Globalisation and Education Reforms: Creating Effective Learning Environments*; Springer International Publishing: Cham, Switzerland, 2021; pp. 35–50.
26. Kolb, D.A. *Experiential Learning: Experience as the Source of Learning and Development*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1984.
27. Zhou, P. Use Semi-Structured Discussion in Health Economics Seminars. In *Global Approaches to Sustainability Through Learning and Education*; Al-Sartawi, A., Hussainey, K., Hannoon, A., Hamdan, A., Eds.; IGI Global: Hershey, PA, USA, 2019; pp. 200–206. [CrossRef]
28. Johnson, M.; Meder, M.E. A meta-analysis of technology: Interventions in collegiate economics classes. *J. Econ. Educ.* **2021**, *52*, 1–16. [CrossRef]
29. Al-Samarraie, H.; Shamsuddin, A.; Alzahrani, A.I. A flipped classroom model in higher education: A review of the evidence across disciplines. *Educ. Technol. Res. Dev.* **2020**, *68*, 1017–1051. [CrossRef]
30. Kay, R.H.; LeSage, A. Examining the benefits and challenges of using audience response systems: A review of the literature. *Comput. Educ.* **2009**, *53*, 819–827. [CrossRef]
31. George, M.L. Effective Teaching and Examination Strategies for Undergraduate Learning during COVID-19 School Restrictions. *J. Educ. Technol. Syst.* **2020**, *49*, 23–48. [CrossRef]

32. Maia, D.; Andrade, R.; Afonso, J.; Costa, P.; Valente, C.; Espregueira-Mendes, J. Academic Performance and Perceptions of Undergraduate Medical Students in Case-Based Learning Compared to Other Teaching Strategies: A Systematic Review with Meta-Analysis. *Educ. Sci.* **2023**, *13*, 238. [CrossRef]
33. Chuang, S. The Applications of Constructivist Learning Theory and Social Learning Theory on Adult Continuous Development. *Perform. Improv.* **2021**, *60*, 6–14. [CrossRef]
34. Gupta, M.; Pandey, J. Impact of Student Engagement on Affective Learning: Evidence from a Large Indian University. *Curr. Psychol.* **2018**, *37*, 414–421. [CrossRef]
35. Ang, C.-S.; Lee, K.-F.; Dipolog-Ubanan, G.F. Determinants of first-year student identity and satisfaction in higher education: A quantitative case study. *SAGE Open* **2019**, *9*, 2158244019846689. [CrossRef]
36. Ryan, M.P.; Glenn, P.A. Increasing One-Year Retention Rates by Focusing on Academic Competence: An Empirical Odyssey. *J. Coll. Stud. Retent. Res. Theory Pract.* **2002**, *4*, 297–324. [CrossRef]
37. Waibel, S.; Rüger, H.; Ette, A.; Sauer, L. Career consequences of transnational educational mobility: A systematic literature review. *Educ. Res. Rev.* **2017**, *20*, 81–98. [CrossRef]
38. Berk, R.A. Survey of 12 strategies to measure teaching effectiveness. *Int. J. Teach. Learn. High. Educ.* **2005**, *17*, 48–62.
39. Skinner, B.F. Selection by Consequences. *Science* **1981**, *213*, 501–504. [CrossRef] [PubMed]
40. Piaget, J. *The Language and Thought of the Child*; Routledge & Kegan: London, UK, 1926.
41. Vygotsky, L. *Thought and Language*; MIT Press: Cambridge, MA, USA, 1962.

applied system innovation

MDPI

Article

Identifying Strengths and Weaknesses in Mobile Education: A Gender-Informed Self-Assessment of Teachers' Use of Mobile Devices

Judith Balanyà Rebollo * and Janaina Minelli De Oliveira

Department of Pedagogy, Rovira i Virgili University, 43002 Tarragona, Spain; janaina.oliveira@urv.cat
* Correspondence: judith.balanya@urv.cat

Abstract: Mobile devices have the potential to transform education and society. Promoting mobile learning and enhancing teachers' digital and entrepreneurial skills are essential in achieving this goal. This study analyses the conditions under which the use of mobile technology can support teachers in the design, implementation, and evaluation of teaching and learning processes. Data were collected using a quantitative method based on a self-assessment instrument (Cronbach's alpha = 1.0046). A total of 327 educators filled out the survey, which included 67 items scored on a Likert scale. The self-assessment tool provided participants with feedback on their mobile device use for educational purposes and suggestions for improvement. The results indicate that the median score of the teachers was 7, which is regarded as satisfactory, with a gender gap of 3.5 points. In addition, three out of seven improvement dimensions were identified: technology learning spaces (54.74%), assessment (57.65%), and design activities (59.26%). In conclusion, the study enabled us to stratify and analyse teachers' pedagogical perceptions of mobile learning and the significance of inference in certain training areas.

Keywords: educational technology; mobile learning; teaching/learning strategies; teacher evaluation

Citation: Balanyà Rebollo, J.; De Oliveira, J.M. Identifying Strengths and Weaknesses in Mobile Education: A Gender-Informed Self-Assessment of Teachers' Use of Mobile Devices. *Appl. Syst. Innov.* **2024**, *7*, 31. https://doi.org/10.3390/asi7020031

Academic Editor: Chun-Yen Chang

Received: 3 March 2024
Revised: 30 March 2024
Accepted: 3 April 2024
Published: 16 April 2024

1. Introduction

Faced with the current educational landscape, marked by rapid technological evolution, mobile devices have consolidated as potential tools to enrich teaching and learning processes [1]. Furthermore, it is essential to recognise that contemporary education has highlighted the relevance of ICT through international reports such as the Education 2030 Agenda, Sustainable Development Goal 4, and the 2017 Quingdao Declaration. Under the Education 2030 Agenda, it is essential to bring information and communication technologies together to strengthen education systems, disseminate knowledge, facilitate access to information, and promote high-quality learning. However, despite the growing presence of mobile devices in society and in education, there is an imperative need to provide teachers with clear strategies and concrete guidelines for the effective pedagogical integration of these devices [2–4].

This research aims to analyse the conditions under which the use of mobile technology can support teachers in the design, implementation, and evaluation of teaching and learning processes. Specifically, it seeks to answer the following question of the investigation: What is the level of theoretical and practical knowledge of mobile learning perceived by the teachers themselves?

To address this issue, a self-assessment tool was developed to help teachers assess their proficiency in designing mobile device activities, focusing on the identified factors that promote their integration in the classroom: content, methodological strategies, activities, evaluation, mobile resources, technological learning spaces, and the teacher [5–9]. In this study, we present results aimed at examining and evaluating the conditions under which the use of mobile devices can support teachers in the design, implementation, and evaluation of teaching and learning processes through the self-assessment tool and analyse their use from a gender perspective.

In this framework, and to carry out the different actions of research, a collaboration was established with the teaching innovation project "PlaMòbils.edu" (Edu/1464/2019, May 27), promoted by the General Directorate for Innovation, Digitization, Study Plans, and Languages of the Department of Education of the Government of Catalonia (Spain). This three-year project sought to improve educational success by using mobile devices (these devices can include smartphones, tablets, laptops, e-readers, 'Wearables', and 'Chromebooks', among others) in the classroom.

This research participated in the second and third phases: implementation (2nd year)—finalisation of planning, follow-up, and evaluation of actions—and project finalisation (3rd year)—final evaluation and transfer. Specifically, the actions were carried out in the area of "Methodologies and resources for the improvement of teaching and learning with mobile devices" that was developed in a virtual classroom of the Campus Ágora (Moodle) for training and research actions for teachers of the Department of Education, covering the period from early childhood education to compulsory secondary education, who were enrolled in the project. The described context emphasises the significance of inter-institutional collaboration and the ongoing training of teachers, which are crucial components for ensuring research and collaboration in a real context.

2. Mobile Devices in the Educational Institutions

The use of mobile devices, such as smartphones and tablets, in educational institutions has become increasingly widespread in recent years [10–15]. These devices offer students and teachers access to a wide range of learning resources and tools and can provide a more flexible and engaging learning experience [16,17]. However, the use of mobile devices in the classroom also poses certain challenges, such as the need for schools to have a clear policy outlining their use and setting guidelines for responsible and appropriate behaviour. It is important to create clear policies for the use of mobile devices in the classroom that address teacher training and support, as this can improve the efficiency and effectiveness of using mobile devices in the classroom [18].

An effective policy should include clear rules regarding the use of mobile devices and how they should be used for educational purposes [19]. This research showed that clear policies can make using mobile devices in the classroom more efficient and effective.

The EU has published the DigComp [20] digital competence framework, which has been revised and adapted for the educational field in which it is defined as the digCompEdu, which is the framework that refers to digital teaching competence. In addition, the European Commission's [20] report, "Digital Education at Schools," highlights that many teachers still have limited knowledge on how to effectively integrate technology into teaching. Although the availability of digital resources for teaching is improving, there are still barriers to their use, such as the lack of high-speed internet access and technological devices. The report also addresses important concerns such as privacy and online security.

At a regional level, the Digital Education Plan of Catalonia (Spain) 2022–2023 [21] establishes a strategic framework for the use of information and communication technologies (ICT) in the educational system. It aims to promote digital inclusion and the development of digital competences in students, as well as to improve the training and professionalisation of teachers.

The plan includes measures for the development of technological infrastructure in schools, the integration of ICT in the curriculum, and the training of teachers in the pedagogical use of technology such as mobile devices. It also addresses the importance of security and privacy in the use of the Internet and ICT.

3. Designing Mobile-Learning Activities

This strategy encompasses initiatives for enhancing the technological infrastructure within educational institutions, incorporating information and communication technology

(ICT) into the educational syllabus, and equipping educators with the skills necessary for the effective pedagogical application of technology, specifically through mobile devices. Furthermore, it highlights the critical significance of ensuring security and privacy in the online environment and the use of ICT.

It is important to make mobile-learning activities that are both interactive and meaningful if you want to keep students interested and help them learn. A previous study has identified core elements to consider when designing mobile-learning activities [6]:

1. The content: this involves considering the intricacy of the content, ensuring it is scaffolded appropriately to cater to the diverse learning needs and technological proficiency of all students. By aligning this approach with the Technological Pedagogical Content Knowledge (TPACK) [22] framework, educators can create a balanced integration of technology, pedagogy, and content knowledge. This harmonisation is crucial for developing instructional strategies that effectively leverage technology to enhance teaching and learning, making education more accessible and inclusive.
2. Methodological strategies: innovative instructional strategies, including gamification, project-based learning, and case-based learning, serve as key elements in elevating the engagement and efficacy of mobile-learning activities. Recognising and accommodating the varied learning styles of students is paramount. By crafting activities that are tailored to meet these diverse needs, educators can ensure a more inclusive and dynamic learning experience. This approach not only enhances student motivation but also promotes a deeper understanding of the material, thereby fostering a more enriching educational environment.
3. Activities: engaging in creating activities that promote collaboration and interaction among students is crucial for enhancing engagement and motivation. The exercises should be carefully organised to not only strengthen the application of learned knowledge in real-world situations but also to progress through the many levels of Bloom's Taxonomy [23]. By following this approach, tasks can be structured to gradually push students to higher levels of thinking—first with fundamental remembering and comprehension, advancing to practical use and examination, and concluding with assessment and innovation. This method guarantees a whole educational experience that enables students to analyse content critically, utilise their knowledge in practical situations, and cultivate advanced thinking abilities, ultimately creating a meaningful and significant learning process.
4. Evaluation: a multifaceted approach to evaluation, encompassing both formative and summative assessments, alongside self-reflection and feedback from peers and instructors, facilitates a thorough understanding of the learning outcomes. This comprehensive assessment strategy not only measures the effectiveness of the learning activities but also encourages continuous improvement and adaptation, ensuring that the educational experiences are both impactful and aligned with learning objectives.
5. Technology resources: utilising technological resources like multimedia, simulations, and interactive elements can make mobile-learning activities more engaging and interactive. It is essential to consider the accessibility of these resources and ensure that they are compatible with the technological platforms utilised by students.
6. Technology learning spaces: the design of technology-based learning environments can have a significant effect on the success of mobile-learning activities. It is essential to consider the space's layout, lighting, and acoustics to ensure that it is conducive to learning and that students can interact with technology and each other effectively.
7. Teachers: the success of mobile-learning activities depends heavily on the teachers. They should be trained to utilise the technology effectively and incorporate mobile learning into their pedagogical practises. Teachers should also be able to provide students with support, such as assistance with technology and instruction on how to utilise the materials effectively.

Figure 1 shows the teaching elements of mobile learning that interrelate and constitute a learning context, taking the student into account.

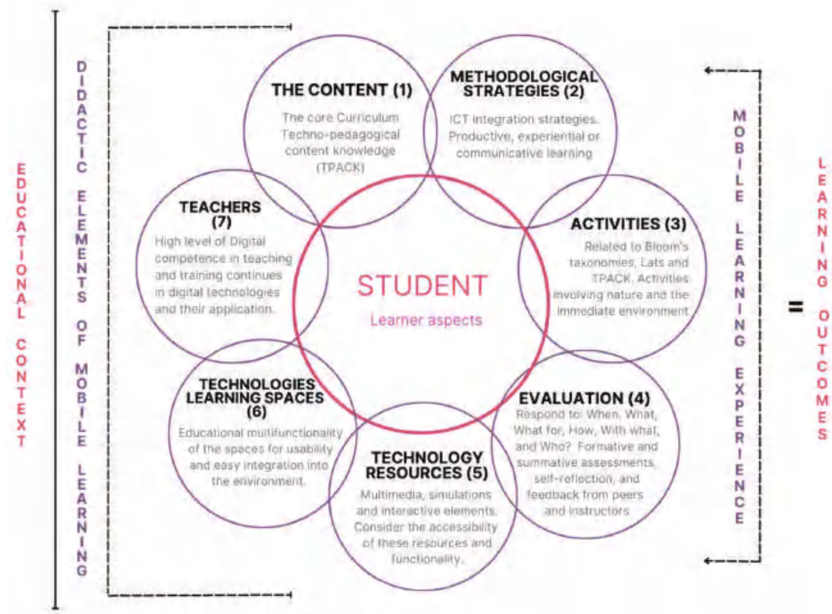


Figure 1. The core aspects of mobile learning.

For the design of mobile-learning activities, it is important that teachers are clear about all the aspects involved. Taking all these elements into account will ensure that their mobile-learning activities are effective and engaging, considering the key factors discussed in this article, such as assessment, accessibility, user experience, etc. However, mobile learning does not come without its challenges. Lack of access to technology and limited connectivity, for instance, can be a hindrance for students in rural or low-income areas. In the mobile-learning environment, the lack of privacy and security of personal information is also a major concern [14,24].

4. Method

This paper aims to analyse the conditions under which the use of mobile technology can support teachers in the design, implementation, and evaluation of teaching and learning processes. Specifically, it seeks to answer the following question of the investigation: What is the level of theoretical and practical knowledge of mobile learning perceived by the teachers themselves?

To address this issue, a self-assessment tool was developed to help teachers assess their proficiency in designing mobile device activities, focusing on the identified factors that promote their integration in the classroom: content, methodological strategies, activities, evaluation, mobile resources, technological learning spaces, and the teacher.

And if so, what is the level of teachers' own perceived theoretical–practical knowledge of mobile learning? This study was carried out under a quantitative methodology with a descriptive and inferential approach, trying to organise, synthesise, and describe all the recollected information from the self-assessment tool. Moreover, the analysis and representation of the research data aim to incorporate a gender perspective. It is currently one of the challenges in terms of gender equality policies to incorporate this perspective in the scientific field [11,12,25]. The consideration of gender perspectives in research is also essential in understanding women and men's experiences and how they interact with educational technology. The Spanish Law 3/2007 of 22 March, for effective equality of women and men, states in Article 20, "Adequacy of Statistics and Studies", that the public powers must systematically include the variable sex in the statistics, surveys, and data collection they carry out [10–13]. This legislation mandates that any research must examine the data from a gender viewpoint, thus guaranteeing that the conclusions are more comprehensive and devoid of gender bias. By following these standards, our aim is to contribute to the advancement of a comprehensive and diverse body of knowledge in the field of educational technology. Furthermore, it is crucial to acknowledge the significance of incorporating the gender variable when designing teacher training programmes and implementing instructional technologies, as these aspects differ between male and female teachers.

This study shows the results of phase 2, which is the second of three phases in the EDR method [26]. This is part of a larger study that looks at how teachers can use mobile devices to help with the design, implementation, and evaluation of E/A processes. Phase 2: In this central phase of the research, the self-assessment, usability, and practicality questionnaire was administered.

4.1. Sample

The study participants were teachers who participated in the pedagogical innovation project called "Pla Mòbils.edu" (Edu/1464/2019, May 27), promoted by the Education Department of Catalonia (Spain) [27]. The group was composed of 60 educational institutions, among them 327 teachers from Catalonia in different educational stages. The table below (Table 1) shows the frequency of participation at each educational stage.

Table 1. Educational stage of the participating teachers.

Educational Stage	**Teachers (*n*)**	**%**
Childhood Education (3–6 years)	20	6.1%
Initial cycle of Primary Education (6–8 years)	33	10.1%
Middle cycle of Primary Education (8–10 years)	37	11.3%
Upper cycle of Primary Education (10–12 years)	68	20.8%
Compulsory Secondary Education (12–16 years)	169	51.7%
Total	327	100%

The majority of teachers, 51.7%, taught in the compulsory secondary education stage (12–16 years), while the smallest number of teachers, 6.1%, taught in childhood education (3–6 years). The other stages had relatively similar representation, with 20.8% in the upper cycle of primary education (10–12 years), 11.3% in the middle cycle of primary education (8–10 years), and 10.1% in the initial cycle of primary education (6–8 years).

4.2. Data Collection Instrument

Data collection was carried out using the validated self-assessment tool [6] and was applied digitally from the project "Pla Mòbils.edu" [27] Virtual Learning Environment (Moodle). Teachers who were part of the project had automatic access to the virtual classroom. The questionnaire was designed with platform tools that allowed us to control a secure environment for both data collection and data processing (Figure 2).

Participants received specific training on digital technologies and mobile devices in education prior to the questionnaire through their enrolment in the educational program "Pla Mòbil.edu", which is based on digital resources, activities, and online seminars.

The questionnaire was divided into two parts: dimension 1 on biodata and dimension 2 on core aspects. Dimension two presented a self-assessment questionnaire consisting of 7 elements (cores aspects): (1) The Content, (2) Methodological Strategies, (3) Activities, (4) Evaluation, (5) Technology Resources, (6) Technology Learning Spaces, and (7) Teachers, with 67 items evaluated using a Likert-type scale of 4 points, 1 being "strongly disagree" and 4 being "strongly agree".

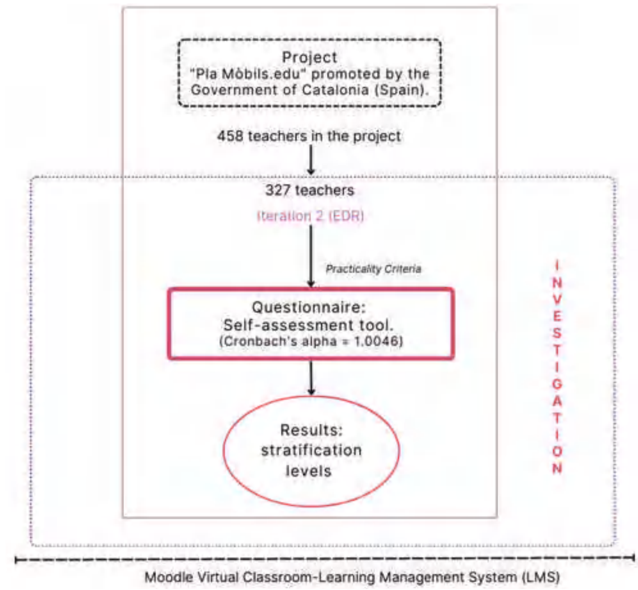


Figure 2. Data collection process.

At the end of the questionnaire, teachers were given a score based on their knowledge on how mobile devices can help them plan, carry out, and evaluate teaching and learning processes (Table 2), as well as feedback based on their scores. This feedback provided them with specific educational resources to improve their pedagogical knowledge.

Table 2. Relation of levels and scores that can be obtained from self-assessment tool.

Type	% Scores	Level
Beginner	0 < 20%	Level 1
Medium	20 < 40%	Level 2
Advanced	40 < 70%	Level 3
Expert	70 < 100%	Level 4

The interaction data from the questionnaire were first analysed with Moodle statistical analysis, and then they were put through the SPPS programme (Statistical Package for the Social Sciences).

5. Results

5.1. Dimension 1: Biodata

The data from the descriptive and inferential analyses of the participants teachers considering the moderating variable of gender are presented below. The goal is to demonstrate the existence or absence of significant differences in order to make future research and gender-specific strategic decisions. This study counted a sample of 327 participants, 34.25% male and 65.75% female. Figure 3 shows the distribution of gender according to teachers' educational stages.

Most of the participants were women, and they outnumbered men in all educational stages. We can also see that in the early stages, women outnumbered men three to one and are thus more visible. On the other hand, male representation increases at the compulsory secondary education stage, but still, women represent 30% more than men.

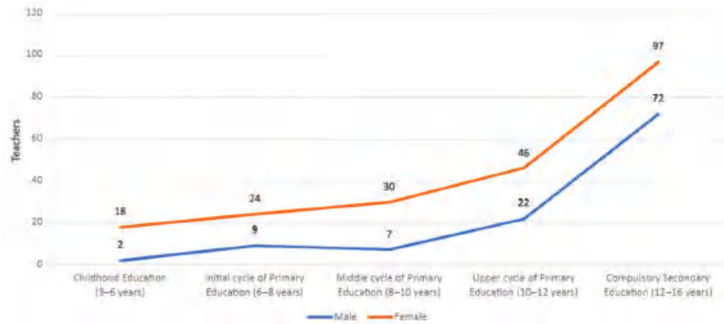


Figure 3. Comparison plot of gender and educational stage.

In Figure 4, we can observe the school subjects taught by the teachers participating in the study. The three most prominent subjects were (1) mathematics (94%), Catalan (87%), and Spanish (76%). Following were "culture and values", "art", and "natural sciences", with less than 35% of remaining subjects technological. It can be said that teachers teach one or more subjects simultaneously, so the percentages are more evenly distributed.

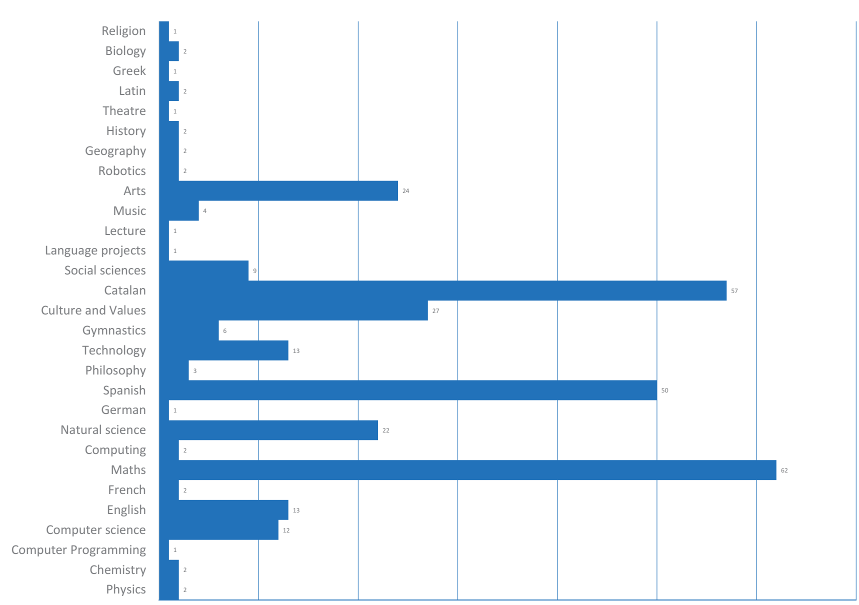


Figure 4. Subjects of participating teachers.

As for the gender distribution of the school subjects (Figure 5), this is notable in subjects dominated more by male (m) than by female (f) teachers: (1) Computing (m = 100% < f = 0%); (2) Computing programming (m = 80% < f = 20%); (3) Physics (m = 72% < f = 28%); and (4) Chemistry (m = 70% < f = 30%). On the other hand, female sex predominated in the following subjects: (1) Art (f = 90% < m = 10%); (2) Catalan, Spanish, and English (f = 80% > m = 20%); (3) Maths (f = 80 < m = 20%); and (4) Culture and Values (f = 80% < m = 20%). Gender representation was balanced in the following subjects: (1) Technology, 50% (male) = 50% (female); and (2) Robotics, 40% (male) > 60% (female).

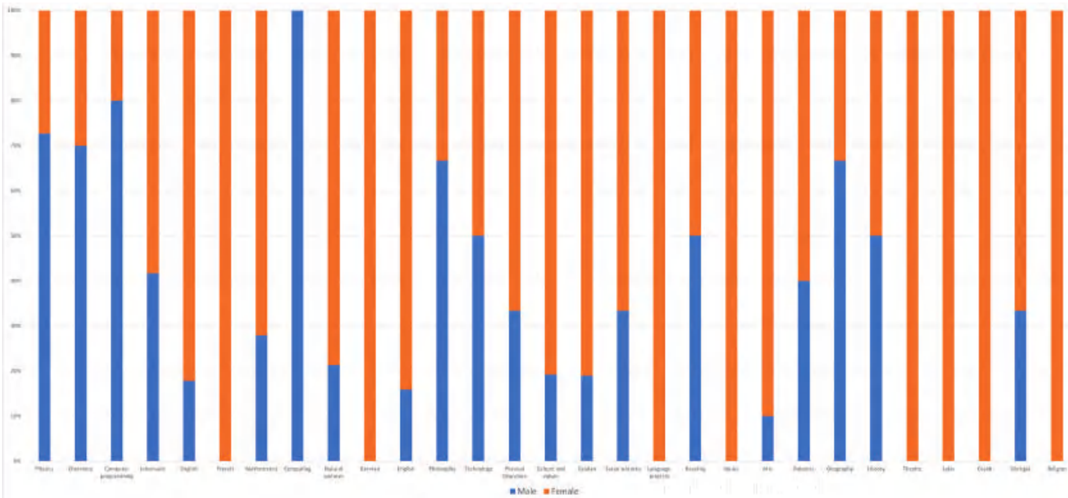


Figure 5. Subjects and gender of participating teachers.

In contrast to the humanities, which typically do not exceed 20%, it is apparent from the subjects of the technological branch that men teach them more frequently than women.

Regarding the frequency of use of mobile devices (Figure 6) as part of their teaching practices, 42.86% of the participants used them occasionally, 30.83% used them frequently, and 25.94% used them very frequently. Only 0.38% of the participants never used mobile devices in their teaching practices.

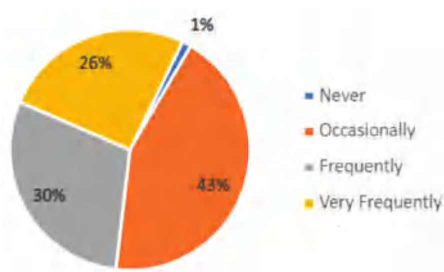


Figure 6. The frequency of using mobile devices in the classroom for educational activities.

Our data indicate that mobile devices are being used in the participants' educational context, not as a highly frequent activity, but on an occasional basis. Figure 7 shows how often mobile devices are used in the classroom.

Figure 7. Frequency of use of mobile devices in the classroom by gender.

Although we can highlight that there is a higher percentage in "occasionally" (85% female, 31% male) and in "frequently" (68% female, <28% male), this is due to the high female participation, specifically 31.5% more.

5.2. Dimension 2: Content Results

The following are the results of the second dimension, which analyses the level of knowledge and use of mobile devices in education. This dimension of the self-assessment questionnaire comprises seven elements: (1) The Content; (2) Methodological Strategies; (3) Activities; (4) Evaluation; (5) Technology Resources; (6) Technology Spaces; and (7) The Teachers.

The data are presented by comparing the averages according to gender and educational stage moderators to establish whether there are significant differences found from the statistical analysis. The arrangement of the data ranges from the general scores of the tool's results to the specifications of each of the 67 items comprising the set of the seven key elements.

Table 3 shows the general median (M) of the group of teachers (*n* = 327), indicating a medium–advanced level of mobile device use in education. The reliability is also identified in the internal consistency, as the partial averages obtained with the different items are consistent with each other; see 1.004 (α). As for the score and the ratio of the levels, in ascending order, 2.96% obtained the middle level, 62.71% the advanced level, and 34.32% the expert level.

Table 3. Global punctuation of teacher self-assessment.

Teachers	M	Mo	σ2	S	Q1	Q3	g2	α
n = 327	7	7	1.865	1.365	6	8	0.061	1.004

The box diagram (Figure 8) allows us to visualise and compare the distribution and central trend of the numerical values of the scores (levels obtained) of teachers through their quarters. The interquartile range (IQR) is 65–77% relative to the global median.

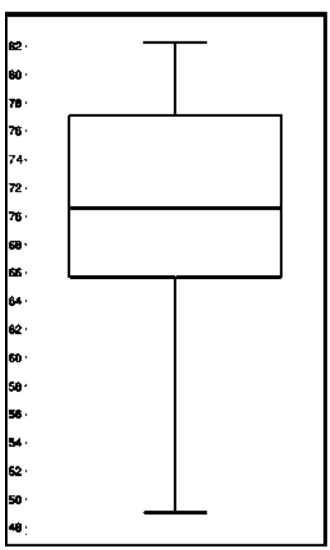


Figure 8. Statistics for the positions of the questionnaire elements.

Table 4 shows the scores obtained in the questionnaire for the seven elements, as can be observed. The Cronbach's Alpha of each element is above 80%, which means that the self-assessment questionnaire blocks have a high level of reliability and, therefore, indicate that the scores received can be used to rank teachers by levels in a consistent manner. Because there are more data values near the average and fewer values in the queues, negative kurtosis is placticurtic.

Table 4. Analysis and ratings of the seven elements of the self-assessment questionnaire.

Teachers	M	σ2	FI	σ	DI	Q3DE
(1) Content	7.75	3.803	78.78%	17.42%	53.40%	62.74%
(2) Methodical Strategies	7.42	4.442	69.80%	19.74%	64.68%	71.10%
(3) Activities	6.83	4.939	66.74%	22.27%	69.63%	75.04%
(4) Evaluation	6.40	5.277	63.41%	22.24%	65.46%	71.56%
(5) Mobile Resources	7.44	4.357	73.52%	21.52%	62.26%	69.48%
(6) Technological Learning Spaces	6.85	5.285	66.31%	23.47%	65.11%	70.25%
(7) Teachers	7.48	4.136	74.78%	20.92%	62.56%	69.78%

In relation to the seven core elements, it can be observed (Table 4) that a Facility Index (FI) between 66 and 80% is fairly easy, showing that teachers have a good understanding of the key ideas that are at the heart of the questions. The Discrimination Index (DI) is greater than 50%, indicating a strong correlation between the scores of the seven elements and the overall test scores. The items with a higher score are as follows: (1) Content 7.75 (M), (2) Teachers 7.48 (7.48), and (3) Mobile Resources 7.44 (M). The elements that show the lowest scores, meaning that teachers have had a more negative perspective on their educational capacity, are as follows: (1) Evaluation 6.40 (M); (2) Activities 6.83 (M) and (3) Technological Learning spaces 6.85 (M). These three elements in turn have variance (V) values that are closer to the median, making them more representative. The dispersion (SD) of the scores with respect to the median is around 20%, a low value, indicating that the answers are close to the median, and in the same way, the discriminative efficiency (DE) (+20%) indicates that the responses of teachers are within the pattern.

In relation to the general scores obtained in Figure 9, the comparison between the educational stages and the median of teachers is shown.

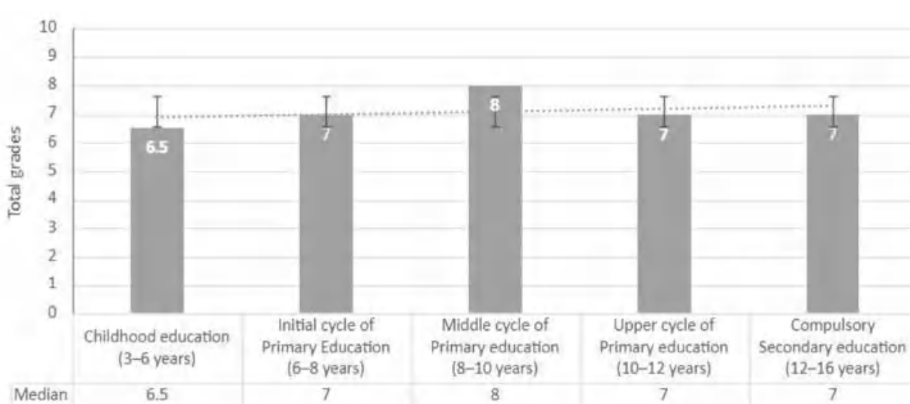


Figure 9. Plot of total grades and educational stages.

The stage with the highest score, meaning that teachers have a positive self-perception of their knowledge, is the middle cycle of primary education (8–10 years) with a median of 8, and the educational level with lowest score is childhood education (3–6 years) with a median of 6.5.

In the following figure (Figure 10), the averages are represented according to the moderating variables of gender and educational stage. As can be seen in the children's educational stages in relation to gender, we can observe that the score obtained is higher in male teachers than in female teachers, even though this is a stage in which most participants are women (6 > 8).

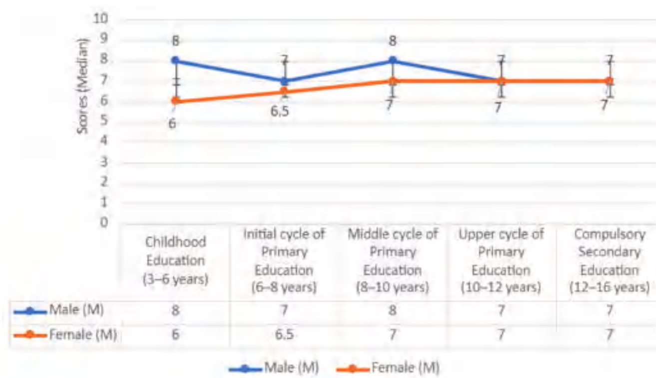


Figure 10. Comparison plot of gender and educational stage scores.

Score and gender differences are not shown during upper education stages, but there is a gender difference during childhood and early primary education stages. Even though male teachers are a minority in the sample and specifically in the childhood educational stages, their median is two points higher: 6 > 8 in childhood education, 6.5 > 7 in the initial cycle of primary school, and 7 > 8 in the middle cycle. Males have a stronger self-perception of their competence in using mobile devices (educational technology) in the classroom than female teachers (relative to the obtained medians). As mentioned above, the sample is made up of 31.5% more women than men, and the declaration of the frequency of use is higher in women (+40%) than in men.

The data collected from each element and items evaluated by teachers are then presented using descriptive analysis. These have enabled us to identify and correlate higher and lower values with various levels of use.

Element 1, "the Content", evaluates and refers to the techno-pedagogical knowledge of the content (TPACK). As can be seen in Figure 11, an advanced level (Level 3) has been obtained with an average of 7.75 (M) and a discriminative efficacy index of 53.40% (FI).

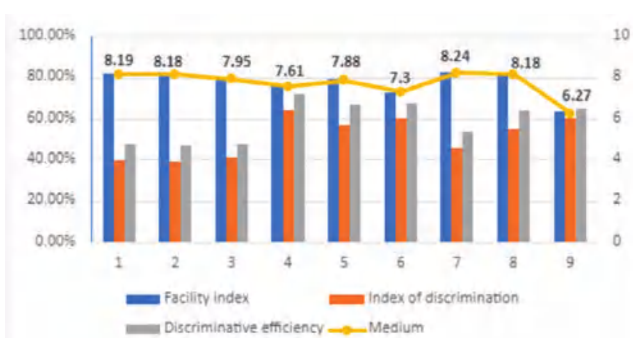


Figure 11. Descriptive results of element 1: content.

The highest and lowest scores were:

- Item 7: I search and select information or open educational resources that best adapt to the educational needs of my students. (M = 8.24, FI = 80%).

- Item 8: I reuse, improve, or create new learning resources that adapt to the educational needs of the students. (M = 8.18; FI < 80%).
- Item 9: I make open access educational resources I've created available to the educational community. (M = 6.27; DE < 60%).
- Item 6: I contrast similar educational experiences that work with mobile devices. (M = 7.3; DE = 65%).

Following the analysis of element 2 (Figure 12), we can see that methodological strategies such as the support of students help in the process of appropriation of knowledge by students through productive, experiential, or communicative learning activities. In this section, teachers obtained an advanced level (level 3) with an average of 7.42 (M) and a discriminative efficiency of 71.10% (DE).

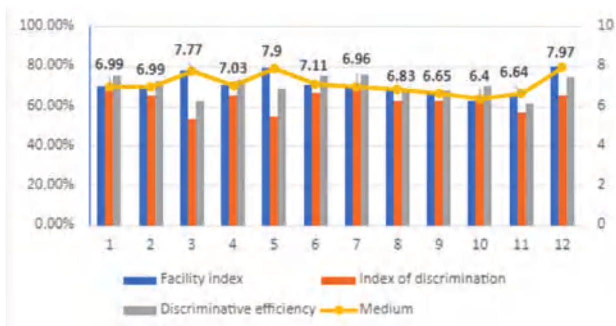


Figure 12. Descriptive results of element 2: Methodological Strategies.

The highest and lowest scores were:

- Item 12: Proposal for incorporating multimedia into methodological strategies for presenting and teaching content. (M = 7.94, FI < 75%).
- Item 5: I apply methodologies that promote students' key competencies. (M = 7.9; FI = 70%).
- Item 10: I apply the Game-based Learning methodology, with the help of mobile devices. (M = 6.4; DE < 65%).
- Item 9: I know that mobile devices can be used to make the Seamless Learning method even better. (M = 6.65; DE < 65%).

Element 3, "Activities" (Figure 13), on the other hand, is concerned with the process of designing activities that are aligned with taxonomies: Bloom, LATs, the TPACK model [28], and activities integrated with the environment [29]. The level obtained corresponds to an average level (level 2) with a median of 6.8 (M) and a discriminatory effectiveness index of 69.63% (DE).

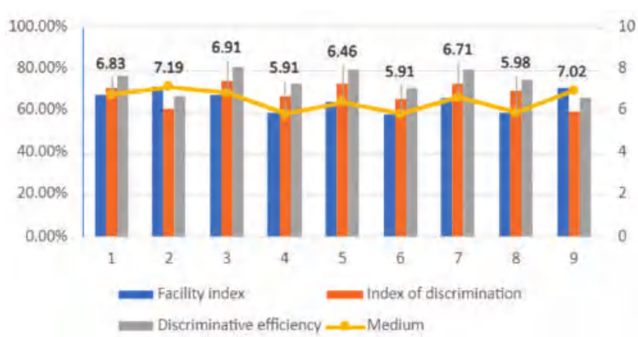


Figure 13. Descriptive results of element 3: Activities.

The highest and lowest scores were:

- Item 2: I differentiate and diversify the activities I propose with mobile devices. (M = 7.19; FI < 75%).
- Item 9: I apply methodologies that promote students' key competencies. (M = 7.2; FI < 70%).
- Item 4: Propose activities related to the taxonomies: 1. cognitive mastery, 2. procedural mastery and 3. attitudinal mastery in educational interventions carried out with students. Attitudinal mastery in educational interventions designed with mobile devices. (M = 5.9; DE < 60%).
- Item 6: Propose activities aimed at projecting, calculating, and reconstructing using mobile devices. (M = 5.91; DE < 60%).

The fourth element, "Evaluation," talks about the steps and questions that need to be thought about when evaluating activities that use mobile devices. In Figure 14, an average level (level 2) with a median of 6.40 (M) and an index of discriminative effectiveness of 65.46% (DE) is shown.

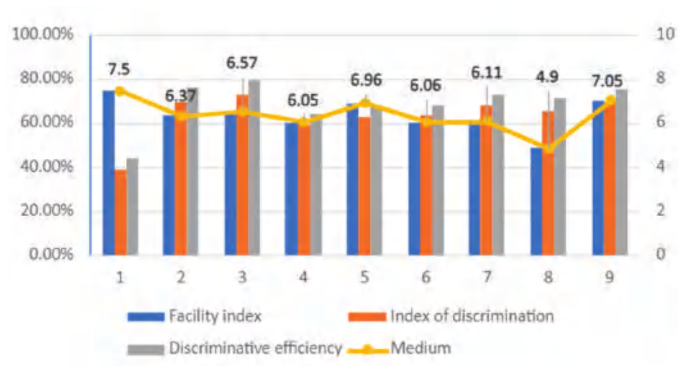


Figure 14. Descriptive results of element 4: Evaluation.

The highest and lowest scores were:

- Item 1: I develop an assessment that answers the following questions: who? (Teachers, students, etc.): What? (Didactic objectives), How? (Quantitatively or qualitatively), When? (Initial, continuous, final, or deferred), Why? (Diagnostic, formative, or summative evaluation), With what? (Information collection instruments) (M = 7.5; DE > 45%).
- Item 9: I will use a variety of evaluation tools in my mobile media activities: interviews, questionnaires, focus groups, observation diaries, reports, projects, checklists, e-portfolios, surveys, individual tests, or observation scales. (M = 7.5; DE = 70%).
- Item 8: I know how to assess students from the 360-degree assessment or integral assessment in the educational interventions that I design with mobile devices. (M = 4.9; DE < 70%).
- Item 4: I have an assessment that contemplates the summative aim of the educational interventions that I design with mobile devices. (M = 6.5; DE < 60%).

Regarding element 5, "Technology Resources," this reflects how technology should be used both in its educational functionality and in its structure [1]. As can be seen in Figure 15, an advanced level (Level 3) has been obtained with an average of 7.44 (M) and a discriminative efficacy index of 62.26% (DE).

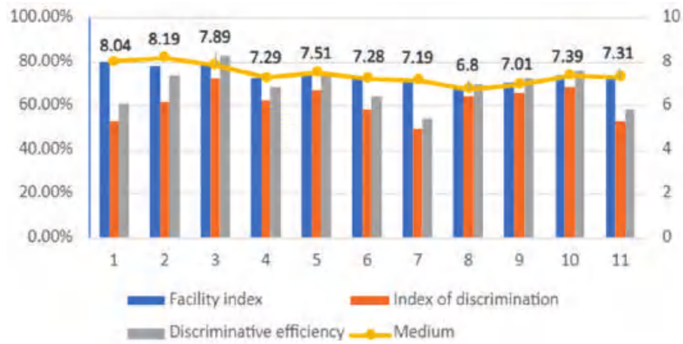


Figure 15. Descriptive results of element 5: Technology Resources.

The highest and lowest scores were:

- Item 2: I know the main characteristics of the mobile applications that I use, and I check their functionality beforehand. (M = 8.19; DE < 70%).
- Item 1: Know the main characteristics of mobile devices. (M = 8.04; DE < 60%).
- Item 8: When I plan learning activities with mobile devices for students, I use both online and offline devices and apps. (M = 6.8; DE < 65%).
- Item 9: I try to establish accessibility criteria in the selection of mobile devices and applications for learners to use. (M = 7.01; DE < 70%).

Element 6, "Learning Space Technologies" (Figure 16), deals with mobile devices, and contemplates their usability and facility integration in the environment so that they do not create protagonism or, at the same time, hinder the educational multifunction of spaces. Teachers here obtained an average level (level 2) with a median of 6.85 (M) and a discriminatory effectiveness index of 70.25% (DE).

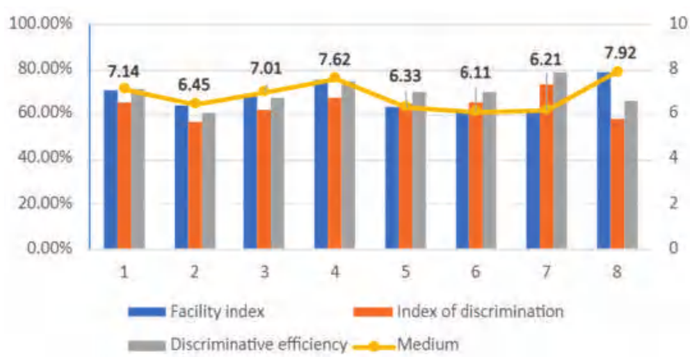


Figure 16. Descriptive results of element 6: Learning Space Technologies.

The highest and lowest scores were:

- Item 8: I consider that spaces with mobile devices must be accessible and facilitate the integration of students from diverse backgrounds. (M = 7.92, DE < 60%).
- Item 4: I take into account the distribution of mobile devices in the classroom for autonomous and group use by the students. (M = 7.62, DE < 70%).
- Item 6: I can create environments in my classroom using mobile devices to create experiences and pique students' interest. (M = 6.11; DE < 65%).
- Item 7: I design spaces with mobile technologies that promote the self-management of information, planning, and reflection on learning by students. (M = 6.21, DE < 75%).

Concerning element 7, "Teachers," the evaluative items assess digital teaching competence and how teachers are trained to use mobile devices in the best way for each situation [13]. As can be seen in Figure 17, teachers obtained an advanced level (Level 3) with an average of 7.48 (M) and a discriminative efficacy index of 69.78% (DE).

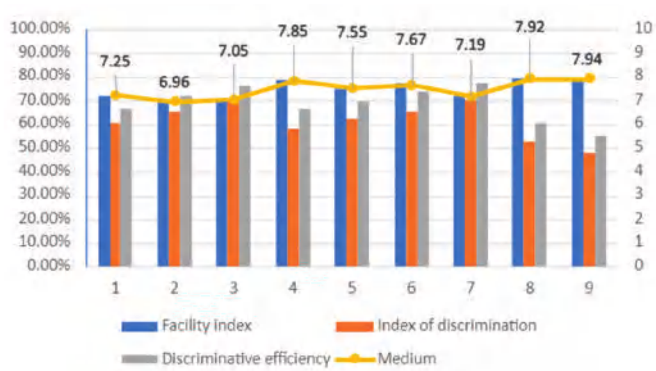


Figure 17. Descriptive results of element 7: Teachers.

The highest and lowest scores were:

- Item 9: I have a high level of teamwork competence with other teaching colleagues at the school. (M = 7.94; DE > 55%).
- Item 8: I believe that I am able to participate in a safe and civic way using my digital identity. (M = 7.92; DE > 65%).
- Item 2: I have a good understanding of ICT and how to incorporate it into the learning process. (M = 6.95; DE < 70%).
- Item 3: I am able to propose techno-pedagogical learning activities for the design, transformation, and application of the content. (M = 7.05; DE < 75%).

In general, this study underscores teachers' proactivity and commitment to adapting and customising educational resources to meet the individual needs of students, leveraging open and mobile technologies. This willingness reflects a positive trend towards educational innovation and the effective use of technology to enrich learning processes. However, it also emerges from the findings that there is a certain reticence or lack of experience in assessing students through mobile devices, suggesting a potential gap in teacher training in this specific area. Assessment, being a critical component of the educational process, requires strategies and tools adapted to technology-mediated teaching modalities. Less confidence or experience in this aspect could limit teachers' ability to fully integrate mobile devices into the classroom in a way that maximises their pedagogical potential. Therefore, it is imperative that professional development programs and teacher training incorporate specific modules on digital assessment, focusing on methodologies and tools that allow for the effective and adaptive measurement of student performance and understanding in technologically enriched environments.

Lastly, the detailed analysis of technology integration in education, as presented in this study, highlights the complexity of this process and the variety of skills required for its successful implementation. As we continue to explore and better understand these dynamics, maintaining a collaborative and research-based approach to developing strategies that support teachers in all facets of technology-mediated education will be crucial.

6. Discussion

The central aim of this research was to analyse the conditions under which mobile technology can support teachers in the design, implementation, and evaluation of teaching and learning processes. More specifically, our investigation sought to unravel the following research question: "What is the level of theoretical and practical knowledge of mobile learning as perceived by the teachers themselves?". To address this, we utilised data derived from a specially designed self-assessment tool, enabling us to capture teachers' perceptions of their use of mobile devices in educational settings.

According to the study's object, teachers use mobile devices in the classroom at a middle–high level (40–70%). Our analysis has revealed insightful details across seven critical dimensions: content, methodical strategies, activities, evaluation, mobile resources, technological learning spaces, and teachers' self-perception.

- Teachers exhibited a strong engagement with content delivery and mobile resources, as evidenced by their high mean scores of 7.75 and 7.44, respectively. The facilitation indices showed that teaching practices effectively facilitated approximately 78.78% and 73.52% of the content and mobile resources domains. These areas also exhibited relatively lower variances, suggesting a consensus among teachers on their comfort and competence in these domains. The robust Cronbach's alpha values (0.855 and 0.897) underscore the reliability of these findings, highlighting content and mobile resources as strengths within the current pedagogical framework.
- Methodical Strategies, Activities, and Evaluation: The domains of Methodical Strategies, Activities, and Evaluation, while still rated positively, indicated areas where teachers may benefit from further support. Mean scores in these categories were slightly lower, with corresponding increases in variance, pointing to a diversity in teacher experiences and comfort levels. Notably, the discriminative efficiency in these areas suggests that these aspects of mobile learning could have a big effect on the quality of teaching, as long as gaps in professional development are filled.
- Technological Learning Spaces and Teachers' Self-Perception: Technological Learning Spaces and Teachers' Self-Perception areas presented a mixed picture. While the facilitation index remained above 66% for both, indicating a generally positive outlook, the higher standard deviations and variances reveal a broader spread of responses. This spread suggests varying levels of confidence and experience among teachers in integrating technology-rich learning environments and reflecting on their own pedagogical practices. The kurtosis values closer to zero indicate a distribution that resembles normality, suggesting diverse experiences that merit further exploration and support.

The findings articulate a clear narrative: teachers are embracing mobile technology in education, recognising its potential to enrich teaching and learning processes. However, the variance in comfort and proficiency levels across different elements underscores the necessity for a nuanced approach to professional development. Specifically, enhancing teachers' skills in assessment, technological learning spaces, and methodical strategies could significantly elevate the quality and efficacy of mobile-learning initiatives.

Reflecting on our research objective and question, it is evident that while teachers possess a foundational theoretical and practical knowledge of mobile learning, as indicated by their perceptions and self-assessment, there remain considerable opportunities for growth. The self-assessment tool has proven to be a valuable instrument in identifying these opportunities, guiding both educators and policymakers in focusing their efforts on areas that promise the greatest impact on pedagogical effectiveness.

These insights should drive policymakers and educational leaders to consider tailored support and professional development in leveraging mobile technologies within the classroom, focusing particularly on areas such as assessment and technological learning spaces where gaps have been identified.

For instance, our findings related to the strong engagement of teachers with content delivery and mobile resources echo the results of studies [30] that emphasise the growing significance of mobile devices as tools for accessing and interacting with educational content. However, our study extends these insights by quantitatively measuring teachers' self-assessment scores and facilitating indices in these domains, thus providing a more granular understanding of their competencies and confidence levels.

Additionally, the research methods, activities, and evaluation parts of our study show areas of mobile learning that need more attention. This is what [31] says about the difficulties teachers face when they try to use mobile technologies effectively in their lessons. Our research adds to this discourse by offering specific data on the variance in teacher experiences and comfort levels, informed by a comprehensive self-assessment tool.

The incorporation of a gender and educational stage perspective also sets our research apart, contributing to a more nuanced understanding of how different demographics engage with mobile learning. This approach resonates with the work of [32], who discusses the importance of considering diverse teacher backgrounds in educational technology research but goes further by providing empirical data on these differences.

Additionally, [33] highlighted the transformative potential of mobile learning, which our study aligns with and expands upon. By quantifying the self-perceived proficiencies of teachers across various pedagogical elements, our research provides insights into the specific supports needed to leverage mobile technology effectively within educational practices.

This research identifies challenges and opportunities related to the adoption and effective integration of mobile learning, which are in line with [34]. However, our analysis, when viewed through the lens of self-assessment and professional development, provides a unique perspective by elaborating on the manifestation and practical solutions for these challenges.

Our study adds to the ongoing discourse on mobile learning within educational contexts by exploring the intricate aspects of how teachers perceive and utilise mobile devices for pedagogical purposes. While previous research has extensively explored the adoption and broad impacts of mobile technologies in education, our investigation seeks to fill a specific gap by focusing on the nuanced pedagogical elements that underpin mobile learning. Through the utilisation of a self-assessment tool, we have been able to gather detailed insights into the areas where teachers feel proficient as well as those where they see potential for further growth. This approach allows us to provide a granular view of mobile learning's implementation challenges and successes, thus offering a complementary perspective to existing literature on the subject.

6.1. Challenges and Opportunities of Self-Assessment Tools of Teacher's Knowledge

Self-assessment tools for teachers in the context of mobile-learning activities present both challenges and opportunities. On one hand, self-assessment tools can help teachers evaluate their own effectiveness and identify areas for improvement. This can lead to more effective and efficient teaching practices [35–38].

The implications of these findings for policy and implementation are profound. By providing teachers with the means to self-assess and reflect on their pedagogical practices, we not only promote personal and professional growth but also pave the way for the development of more dynamic, responsive educational environments. Therefore, it is imperative that educational policymakers and stakeholders consider these insights when formulating strategies for the integration of mobile devices in classrooms. However, the use of self-assessment tools also presents several challenges. One challenge is ensuring the accuracy and reliability of the self-assessment results. Another challenge is the need for teachers to have adequate training and support in using the self-assessment tools. Despite these challenges, self-assessment tools offer several opportunities for teachers in mobile-learning environments. For example, self-assessment can help teachers reflect on their teaching practices and identify areas for improvement.

Policymakers and educational institutions should therefore prioritise the development and dissemination of comprehensive training programs. These programmes should not only focus on the technical aspects of mobile learning but also address pedagogical strategies to enhance teaching and learning outcomes. Self-evaluation is an important tool for personal and professional growth because it lets a person evaluate how well they performed a certain task or activity. Bandura [39] states that self-assessment is an effective strategy for improving self-awareness and self-efficiency. According to the data, out of *n* = 300 female teachers, 25% did not participate, and out of *n* = 158 male teachers, 72.68% did not participate, so we can say that women are more likely than men to self-assess their level of use of mobile devices. Despite this, female participants perceive themselves as less knowledgeable of the pedagogical uses of mobile devices than the male ones. According by Dunning and Kruger [40], women tend to underestimate their abilities, while men tend to overestimate them. This phenomenon is called the Dunning–Kruger effect. In other words, women are more self-critical and have less confidence than men. In addition, even when both sexes have the same outcome, women tend to evaluate their ability in each task lower than men. The term for this phenomenon is "self-assessment bias" [35].

The future action plans should specifically aim to empower female teachers through targeted training and professional development initiatives. Enhancing their confidence and competence in using mobile devices for educational purposes will contribute to a more equitable and effective educational technology landscape [41–43].

The future action plans should focus on capturing female teachers so that they can be trained and trained in the use of mobile devices and thus have a more positive self-perception of their professional abilities. The great participation in the thematic areas of that field is also manifested in the increased participation of the lower educational stages (infantile, primary, and primary), yet there are more teachers in the higher stages. Therefore, training spaces should be provided to teachers to promote the implementation of the use of mobile devices in the initial stages and incorporate training for non-technological subjects.

6.2. Limitations and Future Research Directions

Although our study benefits from a large sample size, it is critical to recognise key limitations that will guide future research. Focusing on teachers in the "Pla Mòbils.edu" project allowed us to analyse mobile device integration in education from a unique and committed standpoint. This unique feature has given us great insights into how these teachers, who are already predisposed to adopting educational technologies, perceive and use mobile devices in their pedagogical practices. Our findings are based on the experiences of instructors in "Pla Mòbils.edu" who are interested in integrating mobile technology into the classroom. This may differ from teachers' perceptions and practices in other educational settings. As a result, one drawback of our study is that the findings may not be applicable to all teachers, particularly those who have not participated in previous projects or who have limited access to resources and training in mobile educational technologies. This highlights the need for additional study with teachers outside of "Pla Mòbils.edu". Expanding our research to include a broader range of educational settings would allow us to determine whether the tendencies observed in our study apply in diverse contexts with varying levels of access to technology and institutional assistance. Future research including instructors from many educational contexts would help us better understand the integration of mobile technology into education by emphasising common obstacles, possibilities, and effective solutions that can be used in a larger range of educational scenarios. Thus, by investigating teachers' adoption and use of mobile devices in a broader range of educational contexts, we can begin to construct a more comprehensive map of mobile technology in education. This will not only enhance our understanding of current dynamics, but will also serve as a strong foundation for the development of policies and practices that support the effective and equitable integration of new technologies at all levels of the educational system.

Currently, society is questioning the educational use of technology, highlighting the need for reflection. Questioning the use of technology in educational environments reflects a maturity point in our society, allowing us to recognise associated risks and highlighting the importance of its proper application through relevant studies. Therefore, monitoring and updating the self-assessment tool in research is crucial in ensuring its long-term relevance. In this sense, the resulting tool must be accessible and free to use for the educational community through the web of research and project 'Pla Mòbils.Edu', aligned with the UNESCO recommendations on Open Science, which promote the sharing of data and results, to facilitate its discussion, feedback, and continuous improvement, allowing the tool to evolve and adapt to the emerging needs of the education field.

As we consider the evolution and expansion of research in the domain of mobile learning, it is crucial to outline potential directions and approaches that could further enrich the field. This study has laid a foundation by exploring the integration of mobile devices from the teacher's perspective, using a self-assessment tool to gauge their comfort and capabilities with these technologies in educational settings. However, the journey does not end here. The horizon of this research encompasses the development of new exploratory phases aimed at broadening the scope of our understanding.

One significant future strategy involves launching a new Educational Design Research (EDR) phase focused specifically on monitoring an entire academic cycle. This phase aims to observe how teachers integrate the self-assessment tool and mobile devices into their pedagogical practices on a day-to-day basis, offering a more nuanced view of the practical challenges and successes encountered. Such an approach will enable a deeper investigation into the impact, adaptability, and long-term viability of mobile-learning tools, fostering a richer comprehension of their effectiveness and potential areas for enhancement within the real-world classroom setting.

In addition, while the insights derived from teachers have been invaluable, a comprehensive assessment of mobile-learning strategies cannot be complete without considering the students' perspectives. It is essential to understand how students perceive, interact with, and impact mobile devices in their learning environments. Their experiences, outcomes, and feedback can provide a complementary lens through which the efficacy of mobile-learning approaches can be more fully evaluated.

Therefore, a pivotal area for future research is to delve into students' viewpoints on mobile learning, examining the perceived benefits, challenges, and overall influence on their educational experiences. Such investigations promise to significantly broaden our understanding of the effectiveness of mobile learning and guide the development of more refined, learner-focused methods for integrating technology into educational frameworks.

By embracing an inclusive approach to research that encompasses both teachers' and students' perspectives, we can ensure that the evolution of mobile-learning environments is informed by a comprehensive understanding of all stakeholders involved. This approach will not only deepen our insight into the dynamics of mobile learning but also aid in crafting policies and practices that genuinely align with the needs and preferences of the wider educational community. By incorporating students' perspectives into our research framework, we ensure that an in-depth appreciation of all parties' experiences shapes the advancement of mobile-learning settings, resulting in more effective, engaging, and inclusive educational experiences.

Author Contributions: Conceptualization, J.B.R. and J.M.D.O.; Methodology, J.B.R. and J.M.D.O.; Validation, J.M.D.O.; Formal analysis, J.B.R.; Data curation, J.B.R.; Writing—original draft, J.B.R.; Writing—review & editing, J.B.R. and J.M.D.O.; Supervision, J.M.D.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: For this study, the collection and processing of data in this research has been carried out taking into account the current regulations on personal data protection. The 2020 'Basic guide for researchers on personal data protection' provided by the Office of Coordination and Advice on Security and Data Protection of the Rovira i Virgili University (URV) has been followed. This guide provides guidance and guidelines to ensure compliance with Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 and the Organic Law 3/2018 of 5 December on the protection of personal data and guarantee of digital rights.

Informed Consent Statement: Informed consent was obtained from the participant involved in the study.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Brown, M.; McCormack, M.; Reeves, J.; Brook, D.C.; Grajek, S.; Alexander, B.; Bali, M.; Bulger, S.; Dark, S.; Engelbert, N.; et al. *2020 EDUCAUSE Horizon Report: Teaching and Learning Edition*; EDUCAUSE: Louisville, CO, USA, 2020.
2. Moya, S.; Camacho, M. Developing a Framework for Mobile Learning Adoption and Sustainable Development. *Technol. Knowl. Learn.* **2023**, *28*, 727–744. [CrossRef]
3. Castañeda, L.; Esteve-Mon, F.M.; Adell, J.; Prestridge, S. International Insights about a Holistic Model of Teaching Competence for a Digital Era: The Digital Teacher Framework Reviewed. *Eur. J. Teach. Educ.* **2022**, *45*, 493–512. [CrossRef]
4. Salcines-Talledo, I.; González-Fernández, N.; Briones, E. The Smartphone as a Pedagogic Tool. Student Profiles as Related to Its Use and Knowledge. *J. New Approaches Educ. Res.* **2020**, *9*, 91–109. [CrossRef]
5. Balanyà Rebollo, J.; Minelli De Oliveira, J. The Crux of Mobile Learning: Key Aspects in Teaching with Mobile Devices. In Proceedings of the 17th International Conference Mobile Learning (ML 2021) and 8th Educational Technologies (ICEduTech 2021), Online, 3–5 March 2021.
6. Balanyà Rebollo, J.; De Oliveira, J.M. Los Elementos Didácticos del Aprendizaje Móvil: Condiciones en Que el Uso de la Tecnología Puede Apoyar los Procesos de Enseñanza y Aprendizaje. *Edutec Rev. Electron. Tecnol. Educ. 2022 80*, 125–219. [CrossRef]
7. Balanyà Rebollo, J.; De Oliveira, J.M. Teachers' Evaluation of the Usability of a Self-Assessment Tool for Mobile Learning Integration in the Classroom. *Educ. Sci.* **2023**, *14*, 1. [CrossRef]
8. Fombona, J.; Pascual, M.A.; Pérez Ferra, M. Analysis of the Educational Impact of M-Learning and Related Scientific Research. *J. New Approaches Educ. Res.* **2020**, *9*, 167. [CrossRef]
9. Kukulska-Hulme, A.; Beirne, E.; Conole, G.; Costello, E.; Coughlan, T.; Ferguson, R.; Fitzgerald, E.; Gaved, M.; Herodotou, C.; Holmes, W.; et al. *Innovating Pedagogy 2020 Exploring New Forms of Teaching, Learning and Assessment, to Guide Educators and Policy Makers*; The Open University: Milton Keynes, UK, 2020. [CrossRef]
10. Aroca-Aroca, M. Concepción de Un Nuevo Paradigma Educativo Desde La Perspectiva de Género. *Cult. Educ. Soc.* **2022**, *13*, 19–40. [CrossRef]
11. Caprile, M.; Valles, N.; Palmen, R. *Guía Para Incorporar la Perspectiva de Género a la Investigación en Salud María del Mar García Calvente*; CIREM Fundación: Barcelona, Spain, 2012. Available online: https://www.ciencia.gob.es/dam/jcr:2ed35333-82b7-492d-afb7-e955d43ac36a/Guia_practica_genero_en_las_investigaciones.pdf (accessed on 25 February 2023).
12. Revuelta, G.; Santamaria Sala, M.; Armengou, C.; Gonzalo, C.; Granero, L.; Richter, A. *Informe Quiral 2018: Mujer, Salud y Comunicación*; Fundación Vila Casas y Centro de Estudios de Ciencia, Comunicación y Sociedad (Universidad Pompeu Fabra): Barcelona, Spain, 2019. Available online: https://www.fundaciovilacasas.com/download-publicacio.php?id=2288 (accessed on 25 July 2022).
13. Sánchez, I. *El Género en la Investigación*; Mujeres y Ciencia del Ministerio de Ciencia e Innovación, Ministerio de Ciencia e Innovación: Madrid, Spain, 2011. [CrossRef]
14. Criollo, S.; Guerrero-Arias, A.; Jaramillo-Alcázar, Á.; Luján-Mora, S. Mobile Learning Technologies for Education: Benefits and Pending Issues. *Appl. Sci.* **2021**, *11*, 4111. [CrossRef]
15. Llopis, M.Á.; Santágueda, M.; Esteve, F.M. Competencia Digital, Actitudes Y Expectativas Hacia Las Tecnologías Digitales. Perfil de Los Futuros Maestros de Primaria. *Rev. Interuniv. Investig. Tecnol. Educ.* **2021**, *11*, 114–130. [CrossRef]
16. Diacopoulos, M.M.; Crompton, H. A Systematic Review of Mobile Learning in Social Studies. *Comput. Educ.* **2020**, *154*, 103911. [CrossRef]
17. Wang, N.; Chen, J.; Tai, M.; Zhang, J. Blended Learning for Chinese University EFL Learners: Learning Environment and Learner Perceptions. *Comput. Assist. Lang. Learn.* **2019**, *34*, 297–323. [CrossRef]
18. Moya, S. Designing a Framework for Mobile Learning Adoption and Sustainable Development. Ph.D. Thesis, Universitat Rovira i Virgili, Tarragona, Spain, 2020. Available online: https://www.tdx.cat/handle/10803/670904?locale-attribute=en (accessed on 3 March 2024).
19. Pelletier, K.; Brown, M.; Brooks, D.C.; McCormack, M.; Reeves, J.; Arbino, N.; Bozkurt, A.; Crawford, S.; Czerniewicz, L.; Gibson, R.; et al. *2021 EDUCAUSE Horizon Report: Teaching and Learning Edition*; EDUCAUSE: Boulder, CO, USA, 2021.
20. Christine, R. *European Framework for the Digital Competence of Educators: DigCompEdu*; Publications Office of the European Union: Luxembourg, 2017. [CrossRef]
21. Direction General of Innovation, Research and Digital Culture of the Government of Catalonia. *Digital Education Plan of Catalonia 2020|2023*; Departament d'Educació: Barcelona, Spain, 2020. Available online: https://repositori.educacio.gencat.cat/bitstream/handle/20.500.12694/309/digital_education_plan_catalonia_2020.pdf?sequence=4&isAllowed=y (accessed on 3 March 2024).
22. Harris, J.; Hofer, M. Technological Pedagogical Content Knowledge (TPACK) in Action: A Descriptive Study of Secondary Teachers' Curriculum-Based, Technology-Related Instructional Planning. *J. Res. Technol. Educ.* **2011**, *43*, 211–229. [CrossRef]
23. Bloom, B.S. Chapter V: Development and Applications of Tests of Educational Achievement. *Rev. Educ. Res.* **1956**, *26*, 72–88. [CrossRef]
24. Fokides, E.; Atsikpasi, P. Tablets in Education. Results from the Initiative ETiE, for Teaching Plants to Primary School Students. *Educ. Inf. Technol.* **2016**, *22*, 2545–2563. [CrossRef]

25. Ariño, M.D.; Tomás, C.; Eguiluz, M.; Samitier, M.L.; Oliveros, T.; Yago, T.; Palacios, G.; Magallón, R. ¿Se Puede Evaluar la Perspectiva de Género en los Proyectos de Investigación? *Gac. Sanit.* **2011**, *25*, 146–150. [CrossRef] [PubMed]
26. Lehrmann, A.L.; Skovbjerg, H.M.; Arnfred, S.J. Design-Based Research as a Research Methodology in Teacher and Social Education—A Scoping Review. *EDeR Educ. Des. Res.* **2022**, *6*, 54. [CrossRef]
27. Diari Oficial de la Generalitat de Catalunya. RESOLUCIÓ EDU/1464/2019, de 27 de Maig, per la Qual es Crea el Programa d'Innovació Pedagògica Mòbils.edu. 2019. Available online: https://agora.xtec.cat/mobilsedu/wp-content/uploads/usu64/2019/06/annex8.pdf (accessed on 5 April 2022).
28. Gamboa Solano, L.; Guevara Mora, M.G.; Mena, Á.; Umaña Mata, A.C. Taxonomía Revisada de Bloom Como Apoyo Para La Redacción de Resultados de Aprendizaje Y El Alineamiento Constructivo. *Innovaciones Educ.* **2023**, *25*, 140–155. [CrossRef]
29. Rikala, J. Designing a Mobile Learning Framework for a Formal Educational Context. Ph.D. Thesis, University of Jyväskylä, Jyväskylä, Finland, 2015. Available online: https://bit.ly/3IBzSW2 (accessed on 3 March 2023).
30. Marimon-Martí, M.; Romeu, T.; Usart, M.; Ojando, E.S. Análisis de la Autopercepción de la Competencia Digital Docente en la Formación Inicial de Maestros y Maestras. *Rev. Investig. Educ.* **2023**, *41*, 51–67. [CrossRef]
31. Sun, J.C. Gaps, Guesswork, and Ghosts Lurking in Technology Integration: Laws and Policies Applicable to Student Privacy. *Br. J. Educ. Technol.* **2023**, *54*, 1604–1618. [CrossRef]
32. Tirado-Morueta, R.; García-Ruíz, R.; Hernando-Gómez, Á.; Contreras-Pulido, P.; Aguaded-Gómez, J.I. The Role of Teacher Support in the Acquisition of Digital Skills Associated with Technology-Based Learning Activities: The Moderation of the Educational Level. *Res. Pract. Technol. Enhanc. Learn.* **2023**, *18*, 010. [CrossRef]
33. Campos, D.; Scherer, R. Digital Gender Gaps in Students' Knowledge, Attitudes and Skills: An Integrative Data Analysis across 32 Countries. *Educ. Inf. Technol.* **2023**, *29*, 655–693. [CrossRef]
34. Tondeur, J.; Van de Velde, S.; Vermeersch, H.; Van Houtte, M. Gender Differences in the ICT Profile of University Students: A Quantitative Analysis. *DiGeSt J. Divers. Gend. Stud.* **2016**, *3*, 57. [CrossRef]
35. Ehrlinger, J.; Johnson, K.; Banner, M.; Dunning, D.; Kruger, J. Why the Unskilled Are Unaware: Further Explorations of (Absent) Self-Insight among the Incompetent. *Organ. Behav. Hum. Decis. Process.* **2008**, *105*, 98–121. [CrossRef] [PubMed]
36. Karaman, P. The Impact of Self-Assessment on Academic Performance: A Meta-Analysis Study. *Int. J. Res. Educ. Sci.* **2021**, *7*, 1151–1166. [CrossRef]
37. Lotero-Echeverri, G. Capacidades de los Docentes Para la Incorporación de Estrategias M-Learning en sus Procesos de Enseñanza y Aprendizaje. Estudio de un Caso Colombiano. *Saber Cienc. Lib.* **2021**, *16*, 220–232. [CrossRef]
38. Sailer, M.; Stadler, M.; Schultz-Pernice, F.; Schöffmann, C.; Paniotova, V.; Husagic, L.; Fischer, F. Technology-Related Teaching Skills and Attitudes: Validation of a Scenario-Based Self-Assessment Instrument for Teachers. *Comput. Hum. Behav.* **2021**, *115*, 106625. [CrossRef]
39. Bandura, A. *Self-Efficacy: The Exercise of Control*; W. H. Freeman & Co., Ltd.: New York, NY, USA, 1997.
40. Kruger, J.; Dunning, D. Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments. *J. Personal. Soc. Psychol.* **1999**, *77*, 1121–1134. [CrossRef]
41. Hinze, A.; Vanderschantz, N.; Timpany, C.; Cunningham, S.J.; Saravani, S.-J.; Wilkinson, C. A Study of Mobile App Use for Teaching and Research in Higher Education. *Technol. Knowl. Learn.* **2023**, *28*, 1271–1299. [CrossRef]
42. Martins, G.; Marcelo, R.; Aparecida, A.; Rosan, R.; Bulla, A. The Use of Mobile Devices in the College Classroom: Project Method and Mobile Technologies in Higher Education. *Rev. Ibero-Am. Estud. Educ.* **2018**, *13*, 500–519. [CrossRef]
43. Riaza, B.G.; Rodríguez, A.I. Students' Perception of the Integration of Mobile Devices as Learning Tools in Pre-Primary and Primary Teacher Training Degrees. *Int. J. Hum. Cap. Inf. Technol. Prof.* **2016**, *7*, 19–35. [CrossRef]

sustainability

MDPI

Article

Research on Evaluation Methods for Sustainable Enrollment Plan Configurations in Chinese Universities Based on Bayesian Networks

Keqin Wang [1,2], Ting Wang [3], Tianyi Wang [3] and Zhiqiang Cai [3,*]

1 Undergraduate Academic Affairs Office, Northwestern Polytechnical University, Xi'an 710072, China; keqinwang@nwpu.edu.cn
2 School of Management, Northwestern Polytechnical University, Xi'an 710072, China
3 Department of Industrial Engineering, Northwestern Polytechnical University, Xi'an 710072, China; wt3377@mail.nwpu.edu.cn (T.W.); wangtianyi@mail.nwpu.edu.cn (T.W.)
* Correspondence: caizhiqiang@nwpu.edu.cn

Abstract: Evaluation methods based on data-driven techniques and artificial intelligence for the sustainable enrollment plan configurations of Chinese universities have become a research hotspot in the field of higher education teaching reform. Enrollment, education, and employment constitute the three key pillars of talent cultivation in universities. However, due to an unclear understanding of their interconnection, universities have yet to establish robust quantitative relationship models, hindering the formation of an evaluation mechanism for sustainable enrollment plan configurations. This study begins by constructing a relevant indicator system and utilizing real enrollment data from a specific university. Through statistical methods such as correlation analysis, it systematically sorts out key variables and identifies seven effective indicators, including average admission score and first-time graduation rate. Subsequently, by using the increase or decrease in enrollment quotas for each major as the experimental target, evaluation models for sustainable enrollment plan configurations aimed at enhancing the advanced education rate are constructed using naïve Bayes networks and tree-augmented Bayesian networks; these are compared with three other classic machine learning methods. The accuracy of these models is evaluated through confusion matrices and receiver operating characteristic curves. Additionally, the Birnbaum importance analysis method is utilized to prioritize remaining variables, ultimately identifying the optimal combination strategy of indicators conducive to the sustainable development of the advanced education rate. The results indicate that the average admission score, transfer rate, and student/teacher ratio are the top 3 prognostic factors affecting the advanced education rate, with the TAN model achieving an accuracy of 96.49%, thus demonstrating good reliability.

Keywords: enrollment plan configurations; sustainable enrollment policies; advanced education rate prediction; importance ranking; indicator combination strategy

Citation: Wang, K.; Wang, T.; Wang, T.; Cai, Z. Research on Evaluation Methods for Sustainable Enrollment Plan Configurations in Chinese Universities Based on Bayesian Networks. *Sustainability* **2024**, *16*, 2998. https://doi.org/10.3390/su16072998

Academic Editors: Savvas A. Chatzichristofis, Jian-Hong Ye, Yung-Wei Hao and Yu-Feng Wu

Received: 16 February 2024
Revised: 28 March 2024
Accepted: 2 April 2024
Published: 3 April 2024

1. Introduction

Enrollment, education, and employment are vital components of talent cultivation in higher education institutions. To achieve sustainable and high-quality development in higher education, establishing a linkagefeedback mechanism is crucial. Only through the mutual promotion of these three components can a virtuous "closed loop" of development be formed. The enrollment work of a university serves as the starting point of the entire talent cultivation system. High-quality student intake serves as a solid foundation for the quality of talent cultivation and is also a significant guarantee for improving employment rates. The consensus among many educational professionals is to promote enrollment through education and to promote employment through enrollment [1,2].

An analysis of the literature using "enrollment, education, and employment" as the primary keywords reveals the predominant themes of linked mechanisms. Publications have notably surged since 2018, reflecting higher education's shift towards high-quality development. In recent years, many universities have been proactive, urgently seeking feasible solutions, particularly in the field of enrollment planning, with mathematical modeling methods emerging as a prominent focus of attention. Wang explored a linked mechanism for employment-oriented program settings and enrollment plan allocation, effectively enhancing the rationality of university self-construction efforts [3]. Jiang used a combination of subjective and objective weighting methods to assign weights to indicators for professional early warning mechanisms, scientifically and reasonably determining the expansion, cultivation, maintenance, alert, and elimination of various majors, achieving a process of natural selection among majors [4].

From a review of the literature, it can be observed that researchers usually start with qualitative or quantitative analysis to comprehensively evaluate enrollment, education, and employment, and establish linkage mechanisms based on indicator systems. Although it has received widespread attention, the construction of linkage mechanisms in 90% of Chinese universities remains in the exploratory stage. Less than 10% of universities propose leveraging enrollment plans with application rates and employment rates as primary evaluation indicators, constructing red, yellow, and blue alerts for the qualitative analysis of enrollment programs. Fewer use databases to establish mathematical models, quantifying and extracting indicators of enrollment, education, and employment situations and combining them with the actual situation of the school to assign reasonable weights [5]. However, when constructing enrollment plan warning standards, model parameters are often determined through methods, such as surveys, which are relatively subjective and cannot reflect the level of attention appropriate to each indicator.

There are two major challenges: the first challenge includes data collection, integration, and analysis mechanisms, and the second challenge comprises the evaluation of sustainable enrollment plan configurations based on student and education quality. Currently, data related to these challenges are housed in different departments of universities, sourced from various databases, and exhibit significant heterogeneity, making integration challenging. Furthermore, without establishing a comprehensive indicator system, researchers either face the challenge of dealing with high-dimensional data with numerous noisy indicators or struggle to extract key information due to the limited number of selected indicators. Moreover, the task becomes more challenging as most machine learning or deep learning algorithms, while capable of producing satisfactory predictive results, often fall short in establishing dynamic adjustment mechanisms for enrollment plans. For example, despite researchers attempting to utilize Bayesian classification algorithms and incorporating them into decision tree learning models to construct predictions, they have not recognized the unique node relationships of Bayesian networks for studying the states of indicator combinations. Bayesian networks provide a powerful tool for analyzing and optimizing complex relationships. Therefore, it is possible to explore this technology, establish a network structure with the advanced education rate as the target variable, and discover key influencing features under different indicator combinations, subsequently identifying the optimal combination strategy.

In conclusion, with the goal of achieving sustainable improvement in the advanced education rate, it is imperative to construct linkage feedback mechanisms as well as evaluation methods for sustainable enrollment plan configurations.

2. Literature Review

2.1. Current Research Status of Feedback Mechanisms

Universities are the main actors in formulating enrollment plans, and the linkage between enrollment, education, and employment is indispensable [6]. Anafnova pointed out that the formulation of enrollment plans is a systematic engineering process that requires comprehensive analysis rather than simple quantity adjustments [7]. Wang further revealed that factors such as socio-economic development environment, school positioning, employment quality, and student quality all influence the formulation of enrollment plans, and these internal and external factors are interrelated [8]. Therefore, to achieve the healthy development of the entire system, it is necessary to integrate and coordinate the three core aspects in universities.

The construction of a linked reform system for enrollment, education, and employment was proposed early on, advocating for high-quality university construction, high-quality graduates, a strong teaching faculty, and teaching quality to promote enrollment. There exists a structural contradiction and a mismatch between the talent cultivation in universities and social demand. University graduates are not necessarily in surplus in terms of quantity but suffer from structural imbalances. Several scholars have analyzed that the root cause lies in the failure to establish a sound linkage mechanism for enrollment, education, and employment. These three aspects are interconnected, and any failure in one aspect will affect the entire education process [9,10].

Additionally, predicting the advanced education rate is an important approach to grasp the effectiveness of enrollment plans and the quality of talent cultivation. Addressing this issue, scholars have conducted in-depth analysis and research from initial expert prediction methods to linear prediction based on time series models and grey system models, and the current mainstream non-linear prediction methods based on machine learning such as neural networks and ensemble learning. However, some issues persist. Despite widespread attention to the interconnection between enrollment, education, and employment, research in this area remains limited in depth. The majority of universities in China are still in the exploration stage, with very few attempting to quantify and extract indicators, combine them with actual school weightings, establish models, and construct early warning systems for enrollment in majors [11–13]. The existing literature shows that the laws governing interconnection remain unclear, with most research focusing on theory rather than empirical studies. The majority of research mainly discusses ideas and frameworks, qualitatively elucidating the relationship between the three components, without establishing quantitative relationship models.

The fundamental task of universities is talent cultivation and achieving a dynamic balance between talent supply and demand. However, there is currently a common issue in universities where the enrollment, education, and employment departments work separately. The importance placed on enrollment as the entry point and employment as the exit point is imbalanced, thus failing to form a virtuous cycle. Through an extensive analysis of the literature, it can be observed that the basic steps of this approach, in sequence, can be summarized as follows [14–16]: establishing an enrollment plan decision support system based on the linkage between enrollment, education, and employment; quantifying and extracting indicators; assigning reasonable weights based on the actual situation of the school; establishing a rational model; calculating the comprehensive quantitative scores of enrollment majors; making qualitative evaluations based on quantification; issuing warnings for majors; and, ultimately, assisting in the construction of a linked mechanism for enrollment, education, and employment in universities.

2.2. Current Research Status of Bayesian Networks

Bayesian networks are a type of probabilistic graphical model tool that utilizes graph structures to represent conditional dependencies between variables. Nodes represent random variables, and directed edges represent causal relationships between variables. Probability inference is conducted through the use of joint probability distributions and conditional probability distributions; probability distributions are updated based on known information. Bayesian network technology effectively handles uncertainty and improves model generalization performance by learning dependencies from data. In recent years, significant research progress has been made in fields such as artificial intelligence, data analysis, and bioinformatics, and Bayesian networks have been widely applied in machine learning tasks such as classification and regression.

The application domains of Bayesian networks extend across various industries, encompassing finance, healthcare, and engineering [17–19]. Within the healthcare sector, Bayesian networks find utility in disease prediction, patient risk assessment, and other critical tasks. By amalgamating diverse medical data sources, they can model intricate disease interrelationships, thereby furnishing personalized predictions and diagnostic recommendations for patients. In finance, they are instrumental in risk management and investment decision-making. Through the modeling of market and economic indicators, they facilitate prognostications of future market trends and investment portfolios, aiding investors in devising more efficacious strategies. Within the industrial sphere, they are harnessed for predictive maintenance and production process optimization. By modeling equipment sensor data and production parameters, they enable the early identification of equipment failure risks, consequently enhancing production efficiency.

However, this method is rarely applied in the field of education, especially in the evaluation of sustainable enrollment plan configurations. Bayesian networks provide a powerful tool for analyzing complex relationship networks. For example, by first identifying key variables, a Bayesian network structure for enrollment plans can be established; then, each node can be analyzed in detail to explore the relationship between nodes. Subsequently, statistical studies can be conducted on different combinations to understand the impact of various indicator combinations on the advanced education rate and uncover key features of influential indicators. The advantages of this approach are evident: by probabilistic inference, the uncertainty of enrollment plan configurations can be quantified; Bayesian networks can effectively model the complex relationships involved in enrollment plans, helping to understand the causality between various variables; by uncovering key features of influential indicators, targeted recommendations can be provided to help decision-makers adjust enrollment plans to improve the advanced education rate; models based on Bayesian networks can provide scientific support for educational decision-making, making decisions more data-driven and reliable.

Although Bayesian networks have achieved important results in probabilistic modeling, uncertainty modeling, multi-source information integration, and other aspects in various fields, they still face some challenges such as the computational complexity of complex models, efficiency of parameter learning and inference, scalability of structure learning, integration of domain knowledge, etc. [20]. To improve model performance and efficiency, researchers have been exploring new algorithms and methods, optimizing algorithms for parameter learning and structure learning, and developing approximate inference methods suitable for large-scale datasets.

3. Indicator Identification and Corresponding Datasets

In current practices, enrollment plans are often developed independently, to a certain extent, by the admissions department. Hence, the lack of interoperability among multiple information systems within university organizations makes it challenging to coordinate effectively. The ultimate goal of this study is to explore the extent to which factors in the enrollment, education, and employment stages influence the enrollment rate, thereby improving the efficiency of sustainable enrollment plan configurations. To address the information silos within universities, it is imperative to establish a comprehensive set of indicators. Statistical analysis methods will be employed to identify correlations among these indicators, thus establishing a set of effective and reasonable metrics. Subsequently, integrating and storing data from various departments becomes necessary. Constructing a sample dataset for enrollment is also imperative for incorporation into subsequent modeling processes, thereby enhancing the reliability of the models and achieving coordinated development in terms of scale and efficiency.

3.1. Construction of an Indicator System

Autonomously formulating enrollment plans is beneficial for the operation of each university, but it also poses a challenge. A practically significant indicator consists of an indicator name and a numerical value. Regarding the principles for constructing these indicators, Yang proposed five aspects: correlation, scientificity, directionality, guidance, and information flow [21]. Alghamdi, on the other hand, divided the principles of indicators into systematicness, representativeness, comparability, and operability [22].

Enrollment is the initial stage of talent cultivation in universities, with enrollment quality being a fundamental factor influencing plan formulation. Due to the unique nature of the Chinese college entrance examination system, the average admission score and transfer rate are of primary concern. Cultivation constitutes a critical aspect of talent development within universities. The provision of adequate teaching conditions is essential for fostering talent and conducting scientific research. Additionally, the state of discipline construction and the fulfillment of training plans profoundly impact talent development. Employment represents the final phase of talent cultivation at universities and serves as a crucial testing ground. Notably, the employment rate and advanced education are equally significant components. Through comprehensive consideration and analysis, this study examines the quality of enrollment, education, and employment, resulting in the establishment of a system comprising 3 comprehensive indicators and 14 sub-indicators, as presented in Table 1.

Table 1. Enrollment, education, and employment indicator system.

Indicators	Sub-Indicators	Code	Description
Enrollment Quality	Average Admission Score	Z_1	Average admission scores for major.
	Application Popularity	Z_2	Application popularity reflects the level of interest and demand.
	First Application Rate	Z_3	Individuals apply for a specific major as a primary choice
	Transfer Rate	Z_4	The ratio transferring to other majors.
Education Quality	Student/Teacher Ratio	Z_5	The student/faculty ratio.
	National First-Class Majors	Z_6	Whether it is a national first-class major.
	First-Time Graduation Rate	Z_7	The ratio of graduating on the first attempt to the total enrollments.
	Course Pass Rate	Z_8	The pass rate in academic performance.
	CET-4 Pass Rate	Z_9	The pass rate in the College English Test Band 4 exam. (CET-6 the same)
	Academic Warning Rate	Z_{10}	The percentage of undergraduates who receive academic warnings.
	CET-6 Pass Rate	Z_{11}	The pass rate in the CET 6 exam.
	Competition Winning Rate	Z_{12}	The percentage of individuals receiving awards in competitions.
Employment Quality	Advanced Education Rate	Z_{13}	The rate of students pursuing further education domestically and abroad.
	Employment Rate	Z_{14}	The ratio of employed and further enrolled students to total graduates.

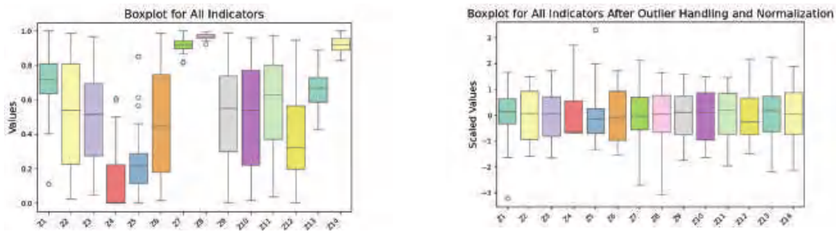
The 14 indicators summarized in this paper are comprehensive and detailed, providing valuable insights for subsequent modeling. However, there may be information overlap and irrelevant associations among the indicators, which could pose challenges to the accuracy and reliability of the model. Therefore, it is necessary to conduct data cleaning and preprocessing, followed by statistical analysis and testing of the sample data, to eliminate irrelevant and disruptive indicators and establish a dataset.

3.2. Construction of the Dataset

After synthesizing the raw data of the 14 indicators provided by the enrollment, education, and employment departments, we obtained sample values for 44 majors at our university over the past decade. Before establishing the model, it is necessary to analyze and clean the dataset to ensure the quality of the modeling dataset.

Cubic spline interpolation fits cubic polynomials between adjacent data points, ensuring smoothness as the function passes through all points with continuous first and second derivatives [23]. That is why we chose this method to fill in missing values in our dataset. Next, considering that Z-scores are calculated based on the population mean and standard deviation, they are more applicable for general outlier handling situations, especially when dealing with non-large sample sizes. We used this method to detect outliers, transforming any outliers using a logarithmic transformation to minimize their impact. Finally, we normalized the data using MinMaxScaler after the transformation.

A comparison of the sample data box plots before and after data preprocessing, as depicted in Figure 1, illustrates that after data cleaning, the data distribution becomes more compact and reasonable. This enhancement demonstrates the effectiveness and comparability of our data preprocessing approach, which is essential for constructing the Bayesian network model.



(**a**) (**b**)

Figure 1. Comparison box plot before and after data preprocessing.

To determine the relationships between different indicators, we can conduct correlation analysis. Using a Pearson or Spearman correlation coefficient, we can evaluate the linear or non-linear correlations between various indicators, which helps identify potential patterns or trends. The correlation heatmap is shown in Figure 2.

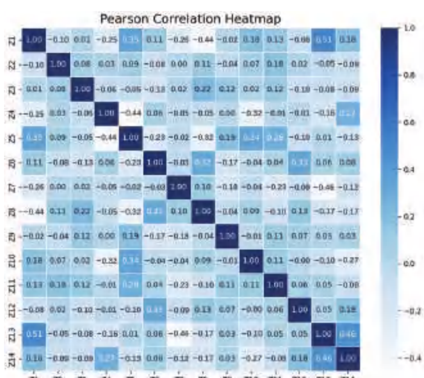


Figure 2. Heatmap of correlations among 14 indicators.

By filtering through a significance level (p-value) less than 0.05, we found significant correlations among Z_1 and Z_5, Z_8, Z_{13}, with a negative correlation with Z_8; we also found significant correlations among the following: Z_4 and Z_5, Z_{10}; Z_5 and Z_8, Z_{10}; Z_6 and Z_8, Z_{12}, Z_7, Z_{13}; Z_{13} and Z_{14}; etc. Taking into account the correlations among the indicators, as well as their correlations with Z_{13}, we ultimately selected indicators with potential experimental value to include in the modeling system. These selected indicators are as follows: Z_1, Z_4, Z_5, Z_7, Z_8, and Z_{14}. This refined set of indicators will help establish a more effective and targeted model. In summary, the six indicators selected through statistical analysis are used as attribute variables, with the advanced education rate as the object variable and the enrollment quota as the decision variable. The variables for further research are shown in Table 2.

The decision variables typically represent factors that can be adjusted and controlled within the model, influencing its output and playing a significant role in decision-making and prediction. Here, decision variable D, representing enrollment quotas, determines if adjustments are needed for specific majors. By controlling enrollment numbers it impacts the study's object variable, the advanced education rate, aiding in evaluating the effects of various enrollment plans.

Table 2. Variables included in the future model.

Variable	Indicators	Code
Attribute Variable	Average Admission Score	X_1
	Transfer Rate	X_2
	Student/Teacher Ratio	X_3
	First Time Graduation Rate	X_4
	Course Pass Rate	X_5
	Employment Rate	X_6
Object Variable	Advanced Education Rate	O
Decision Variable	Enrollment Quota	D

We can categorize the decision variables of enrollment plan quotas into multiple classes. The increase or decrease in enrollment quotas constitutes discrete variables, with each major varying from a reduction of 17 in its quota to an increase of 30. To ensure appropriate categorization—neither excessive nor insufficient—enabling each category to contain a sufficient amount of sample data and avoiding the problem of imbalanced data while ensuring that the decision variables better support sustainable enrollment plan configurations, we divided them into six categories, using a quota of 10 as the boundary. For example, majors with a reduction of up to 10 in their quotas compared to the previous year were categorized as -1. The classification criteria and categories are presented in Table 3, which effectively reflect different enrollment scenarios.

Table 3. Enrollment quota variation classification table.

Classification Status	Degree Interval	Changes in Enrollment Quota Compared to the Previous Year	Frequency	Percentage
−2	(−10, −20]	Decrease by 11 to 20	12	0.02%
−1	(0, −10]	Decrease by 1 to 10	147	0.27%
0	0	No Increase, No Decrease	122	0.23%
1	(0, 10]	Increase by 1 to 10	184	0.34%
2	(10, 20]	Increase by 11 to 20	49	0.09%
3	(20, 30]	Increase by 21 to 30	24	0.05%

Currently, there is a problem of isolated information islands in domestic universities, where internal information systems and application systems are not interconnected. This situation leads to difficulties in accessing unified enrollment, education, and employment data, as these datasets are non-public and specific to each university. Consequently, there is no unified enrollment data platform accessible to the public. Therefore, this study collected data from the enrollment, education, and employment departments of a specific university, established a unified structured database, and built a data service system that integrates hierarchical cooperation, horizontal integration, and logical unity. The detailed dataset statistics from the database are presented in Table 4.

Table 4. Detailed dataset statistics.

Statistic Index	X_1	X_2	X_3	X_4	X_5	X_6	O
count	538	538	538	538	538	538	538
mean	619.2476	0.12285	7.428285	0.969034	0.920058	0.919034	0.653845
std	0.12224	0.095298	0.17136	0.03571	0.02057	0.04562	0.089363
min	568.8829	0	1.269231	0.919995	0.829694	0.817204	0.426087
25%	596.8511	0	4.156353	0.959338	0.888967	0.897759	0.58722
50%	615.5006	0.014651	6.643368	0.969649	0.921862	0.91852	0.671508
75%	624.4629	0.224199	8.50762	0.980844	0.957335	0.945785	0.730256
max	636.69	0.609375	26.2439	0.99384	1	1	0.888889

For all variables in the dataset, their values are continuous. However, the later used Bayesian network can only handle discrete variables. Therefore, we applied the K-means algorithm to partition the values of the variables into two or three ranges. For example, the binary threshold for the object variable was set at 0.638 based on the calculation results. The dataset after discretization is shown in Table 5.

Table 5. Discretization of variables in the dataset.

Variable	Interval	State	Frequency	Percentage
X_1	$\leq$609.246	1	49	9.11%
	609.246 < $X \leq$ 621.251	2	232	43.12%
	>621.251	3	257	47.77%
X_2	$\leq$0.087	1	342	63.57%
	0.087 < $X \leq$ 0.296	2	74	13.75%
	>0.296	3	122	22.68%
X_3	$\leq$6.026	1	232	43.12%
	6.026 < $X \leq$ 14.27	2	257	47.77%
	>14.27	3	49	9.11%
X_4	$\leq$0.961	1	159	29.55%
	0.961 < $X \leq$ 0.977	2	208	38.66%
	>0.977	3	171	31.78%
X_5	$\leq$0.898	1	183	34.01%
	0.898 < $X \leq$ 0.947	2	208	38.66%
	>0.947	3	147	27.32%
X_6	$\leq$0.902	1	147	27.32%
	0.902 < $X \leq$ 0.943	2	171	31.78%
	>0.943	3	220	40.89%
O	$\leq$0.638	1	220	40.89%
	>0.638	2	318	59.11%

4. Methods

4.1. The Technical Roadmap of This Study

The specific research approach of this paper is illustrated in Figure 3.

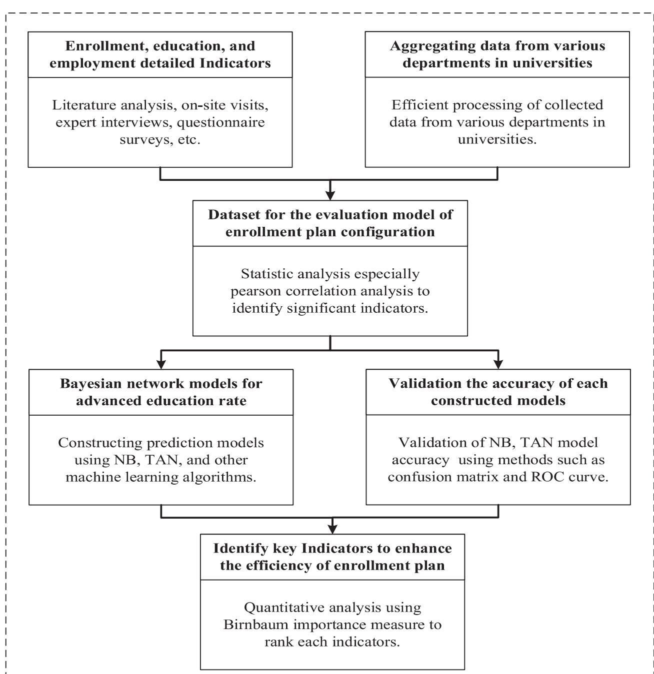


Figure 3. Technical roadmap of this study.

4.2. Statistical Analysis

Statistical analysis is a process of interpreting, summarizing, and inferring data using mathematical and statistical methods, which includes collecting, organizing, and describing data and then using statistical models and inference methods to gain insights into the underlying patterns of the data. Among these methods, Pearson correlation analysis is a statistical method used to measure the linear relationship between two continuous variables, and its formula is as follows [24]:

$$r = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2\left(Y_i - \overline{Y}\right)^2}} \quad (1)$$

where X_i and Y_i are the ith observed values of variables X and Y, respectively, $\overline{X}$ and $\overline{Y}$ are their means, and n is the number of observations. The correlation coefficient r gauges the strength of the linear relationship between two variables. Values closer to 1 or -1 indicate a stronger linear relationship, while those nearer to 0 suggest a weaker one.

4.3. Bayesian Network

The network topology of a Bayesian network is a directed acyclic graph, where nodes represent random variables and arrows indicate relationships. If two nodes are connected, it signifies a causal relationship between them. If one node influences the outcome of another, then a conditional probability value is established. For instance, if node A directly influences the outcome of node B, the directed arc (A, B) is established with an arrow from A to B. The strength of the connection between two nodes is represented by the conditional probability $P(A|B)$. Thus, Bayesian networks classify data samples from a probabilistic perspective, making them an important machine learning method in the field of artificial intelligence. The construction of a network model mainly involves four steps: selecting nodes and variables, determining the network topology, specifying node state probabilities, and determining conditional probability tables between nodes.

The naïve Bayes (NB) classifier is a popular algorithm in Bayesian networks, often used in document classification and spam filtering. It assumes independence among modeling random variables, with each variable related only to one parent node, typically the target prediction variable. The tree-augmented naïve Bayes (TAN) classifier enhances this method by introducing a tree structure to capture dependencies among features, thus relaxing the assumption of attribute independence. Therefore, when dealing with classification problems, especially in cases where correlations exist among features, the TAN classifier method can provide more accurate probability estimates [25,26]. One should suppose that $A_1, A_2 \ldots A_n$ represent n attribute variables, and C represents the target variable. Examples of NB and TAN are illustrated separately in Figure 4.

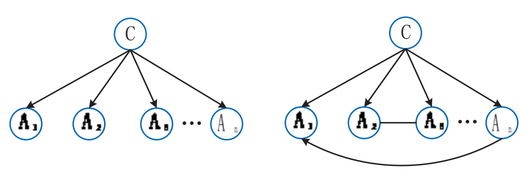


Figure 4. NB and TAN classifiers.

The TAN algorithm effectively balances the complexity and learnability of the model, demonstrating, to some extent, the relationships between nodes. This partially compensates for the limitations of the NB network and aids in the analysis of the relationships among enrollment, education, and employment indicators [27]. The algorithm's description process is outlined in Algorithm 1.

Algorithm 1. The Modeling Process of TAN (TAN Algorithm)

1. Input dataset: obtain the training instance set D.
2. Compute conditional mutual information: $I_{P_D}(A_i; A_j|C)$ *where* $i \neq j$.
3. Build complete undirected graph : where $A_1, A_2, \ldots, A_n$ are attributes associated with respective nodes. Annotate the edge connecting A_i and A_j with the weight of the conditional mutual information: $I_{P_D}(A_i, A_j|\mathrm{D})$.
4. Build a complete undirected maximum-weight spanning tree.
5. Select root attribute and directionalize the tree: set the direction outward.
6. Add class node: add a node labeled as C and introduce an arc from C to A_i.
7. Build TAN model: the resulting directed tree is the TAN model.

4.4. Evaluation Methods

The confusion matrix, also known as the classification matrix, is a fundamental method used to evaluate the credibility of prediction models. The possible results generated by the classifier include true positives (*TP*), false positives (*FP*), false negatives (*FN*), and true negatives (*TN*). *TP* and *TN* describe the number of correctly classified instances, *FP* describes the number of negative samples incorrectly classified as positive, and *FN* describes the number of positive samples incorrectly classified as negative. The diagonal elements of the confusion matrix represent the samples correctly classified by the machine learning classifier. The accuracy of the model (*Acc*) represents the proportion of such samples in the total samples, calculated as follows [28]:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

Recall (*R*), also known as sensitivity, represents the proportion of correctly predicted positive samples among all actual positive samples, as follows [28]:

$$R = \frac{TP}{TP + FN} \tag{3}$$

Precision (*P*) represents the proportion of actually positive samples among all samples predicted as positive. Its calculation formula is as follows [28]:

$$P = \frac{TP}{TP + FP} \tag{4}$$

Two models with low precision and high recall are difficult to compare, and vice versa. To make them comparable, we use the F1 score (*F1*), which ranges between 0 and 1. This metric helps measure both recall and precision simultaneously and represents their harmonic mean, as follows [28]:

$$F1 = \frac{2 * R * P}{R + P} \tag{5}$$

For a machine learning model, it is typically desired to maximize its true positive rate and minimize its false positive rate. However, in real prediction scenarios, both of them usually increase synchronously with the number of positive samples predicted. Therefore, the area under curve (*AUC*) value, which calculates the area under the receiver operating characteristic curve (ROC), is used to evaluate the performance of the model.

4.5. Importance Measures

The importance theory is a branch of reliability mathematics that integrates various cutting-edge and hot research areas such as hazard, sensitivity, risk, and importance. It quantitatively expresses the criticality of each component in a system, clearly indicating the differences in criticality and importance among the components within the system.

There are three commonly used methods, including probability importance, structural importance, and key importance. Probability importance, also known as Birnbaum importance, is based on traditional sensitivity analysis methods. It is obtained from the partial derivative of system reliability with respect to the reliability of individual components, $P_i(t)$, and its formula is typically represented as the following [29]:

$$I^B(i|t) = \frac{\partial h(t)}{\partial p_i(t)}, i = 1, 2, \ldots, n \quad (6)$$

For commonly studied binary systems, importance can be described as the change in the probability of system functioning caused by the change in the state of components. The specific calculation formula is as follows [29]:

$$I(BM)^S_{C_i} = P(S = 1|C_i = 1) - P(S = 1|C_i = 0) \quad (7)$$

where $S = 1$ represents the system being in a functional state, $C_i = 1$ denotes the component being in a working state, and $C_i = 0$ indicates the component being in a failed state. When machine learning algorithms make predictions, they consider the comprehensive states of various indicators, necessitating prioritization among them [30]. This helps identify key predictive indicators, enhancing the credibility and performance of predictive models based on the conclusions and data results.

5. Experimental Results

The undergraduate enrollment rate, which refers to the data of students pursuing further education either domestically or abroad after completing their undergraduate studies, is not only consistent with the top-tier innovative talent training positioning of various universities but also serves as an important indicator for evaluating the outcomes of higher education. Therefore, in the modeling process using the aforementioned 538 cases of discretized data, we selected X_1, X_2, X_3, X_4, X_5, and X_6 as 6 node variables, the advanced education rate (denoted as variable O) as the object variable, and the enrollment quota (denoted as variable D) as the decision variable.

5.1. Naïve Bayes Network Model

Based on the theory of NB networks, we set O as the sole parent node in the model, from which seven unidirectional edges were equally drawn, each pointing towards six attribute variables and one decision variable. We partitioned the sample data into training and test sets randomly to facilitate model training and evaluation. The training set comprised 430 records, while the test set consisted of 108 records, adhering to an 8:2 ratio for the separation of the training and testing sets.

The probabilistic graphical model is depicted in Figure 5. Since it is assumed that the attribute variables are independent of each other, the probability of the predicted outcome depends only on the states of each child node. As the probability distributions of the predictive indicators in the child nodes change, the probability of an increase in the advanced education rate in the target parent node also changes accordingly.

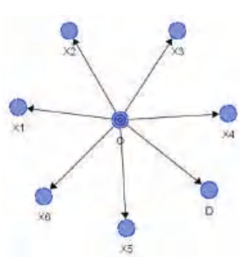


Figure 5. Enrollment plan configuration naïve Bayes model.

Specifically, we employed BayesiaLab software (Version 5.0.1) to apply the NB algorithm for modeling purposes. The relationships between each attribute variable and the target variable were derived from prior knowledge and data analysis. Consequently, we learned the conditional distributions of the edges from the dataset using the software automatically, as illustrated in Figure 6. These distributions pertain to attribute variables, object variables, and decision variables.

O	1	2	3
1	27.778	38.889	33.333
2	0.000	42.308	57.692

X_1

O	1	2	3
1	72.222	0.000	27.778
2	57.692	23.077	19.231

X_2

O	1	2	3
1	44.444	44.444	11.111
2	42.308	50.000	7.692

X_3

O	1	2	3
1	16.667	44.444	38.889
2	34.615	38.462	26.923

X_4

O	1	2	3
1	50.000	38.889	11.111
2	23.077	38.462	38.462

X_5

O	1	2	3
1	16.667	44.444	38.899
2	42.308	38.462	19.231

X_6

O	−2	−1	0	1	2
1	5.556	16.667	11.111	50.000	46.667
2	0.000	34.615	30.769	23.077	11.538

D

1	2
40.9	59.091

O

Figure 6. NB classifier conditional distributions.

After establishing the NB enrollment plan evaluation model, it is crucial to conduct an initial evaluation of the model's performance based on a test dataset to verify its effectiveness. First, we present the confusion matrix of the model. When the decision threshold of the model was set to 0.638, the overall *Acc* was calculated to be 86.54%, as detailed in Table 6.

Table 6. Naïve Bayes model confusion matrix.

Value	1 (167)	2 (371)
1 (220)	157	63
2 (318)	10	308

Next, we plotted the ROC curve of the model, as shown in Figure 7, and calculated the *AUC* value, as well as the overall *Acc*. Finally, we obtained an area under the ROC curve of 94.63%.

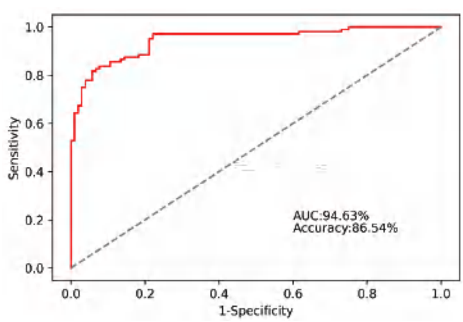


Figure 7. Naïve Bayes classifier ROC curve.

5.2. Tree-Augmented Bayes Network Model

During the modeling process based on the NB classifier, the correlation between various indicators was not considered. However, in actual enrollment planning, many indicators are often not independent, so the correlation between indicators cannot be ignored. The remaining six indicators were used as node variables for modeling, taking into account the mutual information between each node. Subsequently, a prediction model based on the TAN classifier was generated, as shown in Figure 8.

The TAN network still had only one parent node. However, in addition to the single-directional edges extending to six attribute variables and one decision variable, there were certain correlations between some variables. From the graph, it can be observed that there is a significant correlation between X_1 and X_2. This indicates a strong correlation between the two, suggesting that majors with higher average admission scores tend to have lower transfer rates. This is because the arcs generated between nodes using the TAN method cannot only be supported by existing common-sense theories but also uncover some potential dependency relationships between various indicators.

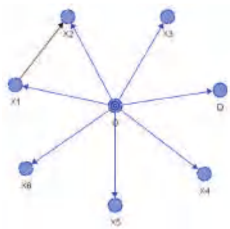


Figure 8. Enrollment plan configuration of TAN model.

The conditional distributions of the edges in the TAN algorithm provided are shown in Figure 9, following the same modeling process as that described in Section 5.1.

X_1

O	1	2	3
1	27.778	38.889	33.333
2	0.000	42.308	57.692

X_4

O	1	2	3
1	16.667	44.444	38.889
2	34.615	38.462	26.923

X_6

O	1	2	3
1	16.667	44.444	38.889
2	42.308	38.462	19.231

X_2

O	X1	1	2	3
1	1	62.500	0.000	37.500
	2	75.000	0.000	25.000
	3	100.000	0.000	0.000
2	1	0.000	54.545	45.455
	2	100.000	0.000	0.000
	3	100.000	0.000	0.000

X_3

O	1	2	3
1	44.444	44.444	11.111
2	42.308	50.000	7.692

X_5

O	1	2	3
1	50.000	38.889	11.111
2	23.077	38.462	38.462

D

O	−2	−1	0	1	2
1	5.556	16.667	11.111	50.000	46.667
2	0.000	34.615	30.769	23.077	11.538

O

1	2
40.9	59.091

Figure 9. TAN classifier conditional distributions.

After establishing the TAN enrollment plan evaluation model, we conducted the evaluation. Firstly, we presented the confusion matrix, still setting the decision threshold to 0.638. At this point, the overall *Acc* was calculated to be 89.42%, as shown in Table 7.

Table 7. TAN model confusion matrix.

Value	1 (205)	2 (333)
1 (220)	184	36
2 (318)	21	297

The plotting of the ROC curve, as shown in Figure 10, ultimately yielded an area under the ROC curve of 96.49%, demonstrating better reliability than the NB method.

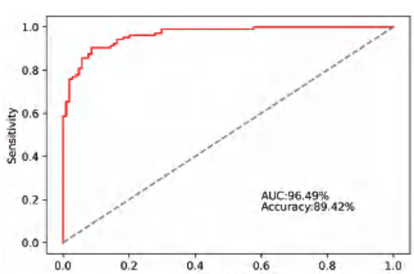


Figure 10. TAN classifier ROC curve.

5.3. Comparison of Classification Model Performance

To validate the effectiveness of Bayesian networks for evaluating sustainable enrollment plans, we conducted modeling experiments using three machine learning methods: linear discriminant analysis (LDA), logistic regression (LR), and support vector machine (SVM). We then performed comparative experiments and evaluated the models using the *P*, *Acc*, *R*, *F1*, and *AUC*. The data types, training sets, and test sets used for modeling were consistent with the Bayesian network modeling process described earlier. The evaluation results of different classifiers, after experiments based on the test set and five-fold cross-validation, are shown in Table 8.

Table 8. Evaluation results of predictions made by different machine learning algorithms.

Model	*P*	*Acc*	*R*	*F1*	*AUC*
NB	0.8750	0.8654	0.8585	0.8632	0.9463
TAN	**0.8750**	**0.8942**	**0.9100**	**0.9001**	**0.9649**
LDA	0.8713	0.8606	0.8462	0.8514	0.9503
LR	0.8738	0.8702	0.8654	0.8672	0.9476
SVM	0.8667	0.8702	0.8750	0.8732	0.9485

For ease of comparison and analysis, we visualized the information from the table, as shown in Figure 11.

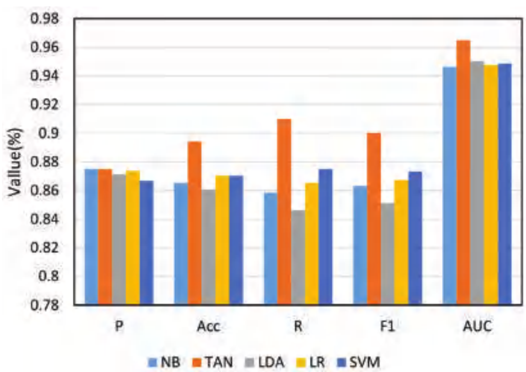


Figure 11. Comparative analysis of predictive performance among various methods.

Combining the model evaluation table and the performance comparison chart, we can observe that there was no significant difference between the precision of several models. However, the recall rate of the TAN algorithm is much higher than that of the other three algorithms, reaching 0.91. Since the *F1* is closely related to *P* and *R*, the *F1* is also highest for the TAN algorithm, followed by that of the SVM, LR, and NB algorithms. In terms of *Acc* and *AUC* value, the TAN algorithm still performs the best. Taking into account all indicators, the TAN algorithm performs the best in *P*, *Acc*, *R*, *F1*, and *AUC* value. The NB, LR, and SVM algorithms have relatively mediocre performance, while LDA performs poorly and is not suitable for this type of data research. Therefore, compared with three machine learning algorithms, namely, LDA, LR, and SVM, the TAN classifier demonstrates certain advantages in prediction, as confirmed by the comprehensive consideration of various performance indicators from different perspectives.

5.4. Calculation of Influence Factor's Importance

Based on the evaluation model, it is possible to lock the initial state by considering different combinations of states of included indicators and then changing another state to observe the probability distribution of the advanced education rate. To quantitatively analyze the importance of each influencing factor, the Bayesian predictive model can be combined with importance theory. By altering the type of state for a specific predictive indicator, the difference in prediction rate can be obtained. Additionally, in many cases, a feature's state is not binary but may be divided into low, medium, and high categories. Based on the composite importance measures method for polymorphic systems, the Birnbaum importance of each predictive indicator can be defined as follows [31]:

$$I(BM)_{C_i}^{S} = \frac{1}{w_i - 1}\sum_{j=1}^{w_i} |P(S = 1|C_i = \mathrm{j}) - P(S = 1)| \tag{8}$$

where S represents the prediction result of X_1, and the remaining indicators are denoted by C. Each variable has a state w, where $P(S = 1)$ represents the prior probability, and $P(S = 1|C_i = \mathrm{j})$ represents the posterior probability. Based on this method, the importance of each influencing factor is shown in Table 9.

Table 9. Birnbaum importance ranking of classification model.

Prognostic Factors	State	Priori Probability	Posterior Probability	MBM	Rank
X_1	0 1 2	0.1136 0.4091 0.4773	0 0.6111 0.7143	1.26785	1
X_2	0 1 2	0.6364 0.1364 0.2273	0.5357 1 1	0.6627	2
X_3	0 1 2	0.4318 0.4773 0.0909	0.5789 0.619 0.5	0.84895	3
X_4	0 1 2	0.3182 0.4091 0.2727	0.7857 0.5556 0.4167	0.879	6
X_5	0 1 2	0.2727 0.4091 0.3182	0.75 0.5556 0.5	0.9028	5
X_6	0 1 2	0.3409 0.3864 0.2727	0.4 0.5882 0.8333	0.91075	4

Through experimentation, we found that the importance of the six scale features—average admission score, transfer rate, student/teacher ratio, employment rate, course pass rate, and first-time graduation rate—decreases sequentially and that all scales are at relatively high levels. It can be inferred that the first three indicators play important roles in the sustainable improvement of the advanced education rate. The reasons behind this are that higher average admission scores, lower transfer rates, and lower student/teacher ratios in various majors represent a higher quality of student intake and education, leading to more apparent educational outcomes and higher chances of successful advanced education. This experimental result is reasonable; there is a strong positive correlation between average admission scores and advanced education rates, while transfer rates and student/teacher ratios are negatively correlated. Therefore, the strategy we propose is to increase the weights of strongly correlated indicators and increase enrollment quotas for majors with high average admission scores, low transfer rates, and low student/teacher ratios. This experimental result underscores the significance of strategic interventions aimed at enhancing the quality of enrollment and education, thereby fostering improved outcomes in advanced education rates.

6. Discussion

Establishing a sustainable enrollment plan evaluation model can comprehensively account for the importance of each stage in enrollment, education, and employment;

streamline the linkage effects between each stage; help schools improve current plans; and complete subsequent plan formulations. Although research on the feedback mechanism has received widespread attention, some issues persist. Previous researchers have attempted Bayesian classification algorithms and integrated them into decision tree models [32], but they did not realize that the unique node relationships of Bayesian networks could demonstrate the states of indicator combinations.

Therefore, we established a Bayesian network structure with the advanced education rate as the object variable and enrollment quotas as the decision variable. By statistically analyzing the states under different indicator combinations, this structure explores the impact of changes in enrollment plans, identifies key influencing features, and identifies the optimal combination strategy conducive to improving the advanced education rate. The Bayesian network model validation confirms that the significance of the six scale features—average admission score, transfer rate, student/teacher ratio, employment rate, course pass rate, and first-time graduation rate—decreases sequentially. The modeling process reflects, to some extent, the real educational environment.

This study primarily focuses on Chinese universities, where the enrollment system, educational structure, and employment market differ from those in other countries. Due to the confidential nature of enrollment data in major Chinese universities, comprehensive data collection for research purposes is challenging. Therefore, we selected the author's affiliated university as the sample. Although the Bayesian network model performed well on this dataset, future efforts should aim to expand data sources. Furthermore, the model itself has limitations such as assuming conditional independence between specific variables, which may not hold true in practice.

However, this does not imply that our model has a regional limitation. Establishing a Bayesian network predictive analysis model allows for a quantitative analysis of factors, facilitating the optimization of current enrollment plans and the formulation of guidance for future plans. Each university can clarify the indicators affecting enrollment plan formulation by considering its national, social, and regional context, as well as its own development status. Then, based on our scheme, key indicators can be selected to establish Bayesian network models, optimize parameters, and propose the optimal combination strategy of indicators aimed at improving the advanced education rate. The broad applicability of the model developed in this study enables its use as a reference for enrollment planning in other universities. Additionally, adjusting parameters based on practical feedback and iteratively optimizing the model further enhances its utility.

7. Conclusions

To analyze the linkage mechanisms and explore how to evaluate sustainable enrollment plan configurations, this study utilized real enrollment data from majors at a university over recent years. Firstly, the challenge of collecting, integrating, and analyzing data from multiple departments, coupled with the absence of a comprehensive indicator system, presents difficulties when handling high-dimensional data or extracting key information. This study addressed this challenge by identifying key variables across three stages and conducting correlation analysis. Seven effective indicators—the student/teacher ratio, transfer rate, average admission score, employment rate, first-time graduation rate, course pass rate, and advanced education rate—were selected.

Secondly, faced with the challenge of evaluating the configuration of enrollment plans based on student and education quality, this study selected the Bayesian network model. The advanced education rate was used as the object variable, and the enrollment quotas were used as the decision variable. This study then explored the mapping relationship from enrollment, education, and employment big data to enrollment plan configuration, evaluating model accuracy through confusion matrix and ROC curve analysis. Comparative experiments were conducted with NB, LDA, LR, and SVM machine learning models, with the TAN model achieving the highest accuracy at 96.49%.

Finally, using the Birnbaum importance analysis method to prioritize the remaining variables, this study explored the impact of changes in enrollment plans, under different indicator combinations, on advanced education rates. The results showed that the average admission score, transfer rate, and student/teacher ratio have extremely significant effects on improving the advanced education rate. In the future, the model can be extended to universities outside of China. Efforts will be made to enhance data collection and analysis mechanisms, refine key indicators for a more accurate evaluation of enrollment plans, further optimize the Bayesian network model, explore more effective parameter optimization methods, and conduct additional empirical research to accumulate practical experience. Strengthening collaborations with international institutions will also be pursued to provide targeted policy recommendations for optimizing enrollment policies, thereby promoting the sustainable development of the higher education system.

Author Contributions: Conceptualization, K.W. and T.W. (Ting Wang); methodology, Z.C.; software, T.W. (Tianyi Wang); validation, Z.C. and K.W.; formal analysis, T.W. (Ting Wang); investigation, Z.C.; resources, K.W.; data curation, K.W.; writing—original draft preparation, K.W.; writing—review and editing, T.W. (Ting Wang); visualization, T.W. (Tianyi Wang); supervision, Z.C.; project administration, K.W.; funding acquisition, T.W. (Tianyi Wang) All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Key Research Project on Undergraduate and Higher Continuing Education Teaching Reform in Shaanxi Province [Grant No. 21BG007]; the Key Research Project on Educational Teaching Reform at Northwestern Polytechnical University [Grant No. 2022JGWG03].

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Thelwall, M. *Word Association Thematic Analysis: A Social Media Text Exploration Strategy*; Synthesis Lectures on Information Concepts, Retrieval, and Services; Springer Nature: Berlin/Heidelberg, Germany, 2021; Volume 13, pp. 1–111.
2. Song, P.C.; Chong, H.Y.; Ong, H.C.; Looi, S.Y. A Model of Bayesian Network Analysis of The Factors Affecting Student's Higher Level Study Decision: The Private Institution Case. *J. Telecommun. Electron. Comput. Eng. (JTEC)* **2016**, *8*, 105–109.
3. Wang, Y. Construction elements and path of practical education model in universities. *Eurasia J. Math. Sci. Technol. Educ.* **2017**, *13*, 6875–6882. [CrossRef] [PubMed]
4. Jiang, Y.; Jin, C.; Yu, H.; Wu, J.; Chen, C.; Yuan, Q.; Huang, W.; Hu, Y.; Xu, Y.; Zhou, Z.; et al. Development and validation of a deep learning CT signature to predict survival and chemotherapy benefit in gastric cancer: A multicenter, retrospective study. *Ann. Surg.* **2021**, *274*, e1153–e1161. [CrossRef] [PubMed]
5. Xiang, B.; Wang, H.; Wang, H. Is There a Surplus of College Graduates in China? Exploring Strategies for Sustainable Employment of College Graduates. *Sustainability* **2023**, *15*, 15540. [CrossRef]
6. Jacobi, M. *College Student Outcomes Assessment: A Talent Development Perspective. ASHE-ERIC Higher Education Report No. 7, 1987*; Association for the Study of Higher Education, Dept. E, One Dupont Circle, Suite 630: Washington, DC, USA, 1987.
7. Anafnova, S. The role of rankings in higher education policy: Coercive and normative isomorphism in Kazakhstani higher education. *Int. J. Educ. Dev.* **2020**, *78*, 102246. [CrossRef] [PubMed]
8. Wang, C.; Zhu, K.; Hedström, P.; Li, Y.; Xu, W. A generic and extensible model for the martensite start temperature incorporating thermodynamic data mining and deep learning framework. *J. Mater. Sci. Technol.* **2022**, *128*, 31–43. [CrossRef]
9. Colbert, B.A. The complex resource-based view: Implications for theory and practice in strategic human resource management. *Acad. Manag. Rev.* **2004**, *29*, 341–358. [CrossRef]
10. Middaugh, M.F. *Planning and Assessment in Higher Education: Demonstrating Institutional Effectiveness*; John Wiley & Sons: Hoboken, NJ, USA, 2011.
11. Raja, R.; Nagasubramani, P.C. Impact of modern technology in education. *J. Appl. Adv. Res.* **2018**, *3*, 33–35. [CrossRef]
12. Altbach, P.G. The many traditions of liberal arts—And their global relevance. *Int. High. Educ.* **2016**, *84*, 21–23. [CrossRef]

13. Tanveer, M.; Verma, S.; Sharma, R.; Goel, T.; Suganthan, P.N. Weighted Kernel Ridge Regression based Randomized Network for Alzheimer's Disease Diagnosis using Susceptibility Weighted Images. In Proceedings of the 2023 International Joint Conference on Neural Networks (IJCNN), Neural Networks (IJCNN), Gold Coast, Australia, 18–23 June 2023; pp. 1–8.
14. Gupta, S.; Garg, S. National Education Policy–2020: Innovations in Higher Education and Quality Assurance. *Theme Aiu Zon. Vice Chancellors' Meet* **2021**, *18*, 14.
15. Zhang, X.; Liu, J.; Zhang, C.; Shao, D.; Cai, Z. Innovation Performance Prediction of University Student Teams Based on Bayesian Networks. *Sustainability* **2023**, *15*, 2335. [CrossRef]
16. Peyser, J.A. Settle for Better: How overpromising undercut the education reform movement, and what to do about it. *Educ. Next* **2023**, *23*, 44–50.
17. Liu, X.; Gao, C.; Li, P. A comparative analysis of support vector machines and extreme learning machines. *Neural Netw.* **2012**, *33*, 58–66. [CrossRef]
18. Faggian, A.; McCann, P. Universities, agglomerations and graduate human capital mobility. *Tijdschr. Voor Econ. En Soc. Geogr.* **2009**, *100*, 210–223. [CrossRef]
19. Wang, H.; Hu, D. Comparison of SVR and LS-SVR for regression. In Proceedings of the 2005 International Conference on Neural Networks and Brain, Beijing, China, 13–15 October 2005; Volume 1, pp. 279–283.
20. Zhuang, Y.; Gan, Z. A machine learning approach to enrollment prediction in Chicago Public School. In Proceedings of the 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, 24–26 November 2017; pp. 194–198.
21. Yang, F. Decision tree algorithm based university graduate employment trend prediction. *Informatica* **2019**, *43*, 573–579. [CrossRef]
22. Alghamdi, J.; Holland, C. A comparative analysis of policies, strategies and programmes for information and communication technology integration in education in the Kingdom of Saudi Arabia and the republic of Ireland. *Educ. Inf. Technol.* **2020**, *25*, 4721–4745. [CrossRef]
23. Huang, H.; Shih, P.C.; Zhu, Y.; Gao, W. An integrated model for medical expense system optimization during diagnosis process based on artificial intelligence algorithm. *J. Comb. Optim.* **2022**, *44*, 2515–2532. [CrossRef] [PubMed]
24. Dixon, W.J.; Massey, F.J., Jr. *Introduction to Statistical Analysis*; McGraw-Hill: New York, NY, USA, 1951.
25. Zhu, M.Q.; Pang, H. Toward a cross-cultural conversation: Liberal arts education for engineers in China and the US. In Proceedings of the ASEE's 123rd Conference & Exposition, New Orleans, LA, USA, 26–29 June 2016.
26. Peterson, P. Liberal education in the global perspective. *Int. High. Educ.* **2011**, *62*, 10–11. [CrossRef]
27. Lee, K.W.; Chung, M. Enhancing the link between higher education and employment. *Int. J. Educ. Dev.* **2015**, *40*, 19–27. [CrossRef]
28. Shalev-Shwartz, S.; Ben-David, S. *Understanding Machine Learning: From Theory to Algorithms*; Cambridge University Press: Cambridge, UK, 2014.
29. Cai, Z.; Si, S.; Dui, H.; Sun, S. Relationship and Changing Analysis of Birnbaum Importance for Different Components with Bayesian Networks. *Qual. Technol. Quant. Manag.* **2013**, *10*, 203–219. [CrossRef]
30. Atuahene-Gima, K.; Murray, J.Y. Exploratory and exploitative learning in new product development: A social capital perspective on new technology ventures in China. *J. Int. Mark.* **2007**, *15*, 1–29. [CrossRef]
31. Miziuła, P.; Navarro, J. Birnbaum importance measure for reliability systems with dependent components. *IEEE Trans. Reliab.* **2019**, *68*, 439–450. [CrossRef]
32. Fang, F. Analysis and Prediction of Higher Vocational Students' Employment Based on Decision Tree Classification Algorithm. In *International Conference on Frontier Computing*; Springer Nature: Singapore, 2021; pp. 973–977.

sustainability

MDPI

Article

Examining the Effects of "Small Private Online Course and Flipped-Classroom"-Based Blended Teaching Strategy on First-Year English-Major Students' Achievements

Luyan Zheng [1,]* and Keok Cheong Lee [2]

[1] School of Foreign Languages, Heze University, Heze 274000, China
[2] Faculty of Education, City University Malaysia, Petaling Jaya 46100, Malaysia; lee.cheong@city.edu.my
* Correspondence: honeynita@163.com

Abstract: Blended teaching, characterized as a combination of online instruction and face-to-face teaching is effective in maintaining high student retention, promoting learners' motivation and saving costs in EFL teaching. However, as low learning efficiency can lead to a reduced performance related to students' blended learning, it is necessary to study the effectiveness of blended teaching. This study aims to examine the effects of the "SPOC and Flipped classroom"-based blended teaching strategy on first-year English-major students' achievements in five English language skills compared with those of the traditional face-to-face classroom teaching strategy. For this research, a quasi-experiment was conducted for one semester, employing an "SPOC and Flipped classroom"-based blended teaching strategy in two undergraduate classes for first-year English-major students. A total of 64 students majoring in English at a Chinese university in Shandong province participated in the quasi-experiment and were divided into the control class (N = 32, (intact group), with 6 men and 26 women) and the experimental class (N = 32, (intact group), with 4 men and 28 women). Intervention was performed in the experimental class, while a typical face-to-face classroom teaching strategy was employed in the control class. Findings of students' overall achievements showed that the blended teaching strategy based on the SPOC and flipped-classroom approach was more effective in improving students' achievements than face-to-face classroom teaching. Findings of students' achievements in each of the five English language skills showed that the "SPOC and Flipped"-based blended teaching strategy was effective in enhancing students' listening, reading, translating, and writing, but was not effective in speaking. Furthermore, students' genders and regional backgrounds were considered as moderating variables of students' achievements, and findings indicated that gender had no significantly positive effect on students' achievements in the blended teaching intervention. However, a significant difference in achievements between the students from urban and rural areas was observed, which indicated that regional background had significantly positive effect on students' achievements in blended teaching. The findings of the research implied that the implementation of well-designed blended learning that uses effective strategies could significantly improve students' achievements in English language skills, but that there would be different results among students of different skill levels and regional backgrounds. Moreover, as the quality of EFL blended teaching improves, students' academic performance will be enhanced.

Keywords: blended teaching strategy; SPOC; flipped classroom; English major; achievements

Citation: Zheng, L.; Lee, K.C. Examining the Effects of "Small Private Online Course and Flipped-Classroom"-Based Blended Teaching Strategy on First-Year English-Major Students' Achievements. *Sustainability* **2023**, *15*, 15349. https://doi.org/10.3390/su152115349

Academic Editors: Savvas A. Chatzichristofis, Jian-Hong Ye, Yung-Wei Hao and Yu-Feng Wu

Received: 10 September 2023
Revised: 6 October 2023
Accepted: 10 October 2023
Published: 27 October 2023

1. Introduction

The term "blended learning" originated in the business world in connection with corporate training [1]. It was later employed in higher education [2] and then in language teaching and learning. Blended teaching has been distinguished from other expressions. Regarding the application of the term to computer/web technology and face-to-face classroom teaching, it is sometimes referred to as "hybrid" or "mixed" [3]. The expression "hybrid teaching", however, refers to "an approach to teaching that not only integrates technology in the teaching process but also combines students who are inside a physical classroom and students from online" [4] (p. 2), which means that not all students are involve in synchronous online and offline learning; instead, optimal methods are taken from both models. This kind of hybrid learning is not the teaching model referred to in this research. "Blended teaching", as addressed in this research, refers to the combination of both online and face-to-face classroom teaching methods, by which all students participate in both learning environments and activities [5]. Face-to-face classroom teaching refers to the traditional teaching model, in which teacher and students carry out the teaching and learning activities synchronously within the same time and space.

Bilgin (2013) points out that "teaching English blending face-to-face teaching with an online LMS (Learning Management System) can be beneficial over solely in-class teaching". Effective learning assessments are crucial to the success of blended learning approaches [6]. Bonk et al. (2005) also suggested that further research and innovation in the blended learning arena would help determine the key contributions, benefits, and impact areas; therefore, it was quite necessary to evaluate whether blended teaching was effective in EFL instruction and should be developed in its future [7]. Blended teaching provides benefits to both teachers and students by facilitating communication and collaboration among students and teachers through social networking, increasing ease of use of course materials [8], decreasing physical class time, creating a student-based learning environment, producing an encouraging learning environment, providing flexible learning times and locations, promoting independent learning skills, and developing individualized course solutions [9].

Although blended teaching could be more advantageous in compensating for the shortages of both traditional face-to-face teaching and purely online teaching, and provide more time, space, and interaction opportunities for teachers and students, the results of some studies show that blended model was merely more effective in student retention, but no significant difference in students' achievements was found [10]. The results implied that blended teaching was not necessarily better than traditional face-to-face teaching in guaranteeing the effectiveness of EFL instruction. Furthermore, the results of a study on blended teaching among college English students [11] showed that the effects of blended teaching for freshmen still required a period of time to be obvious. According to some studies [12], the specific nature of learning faced by English-major students means that first-year English-major students face more difficulties in adapting to blended learning. The practice of blindly promoting blended teaching while ignoring the differences in regional, cultural, and verbal factors among English-major students even could cause "Learned Helplessness under Blended Teaching" [13], which refers to the sense of helplessness experienced by first-year English-major students under the blended teaching model. "Learned Helplessness under Blended Teaching" involves inner barriers, psychological fear, and other states, as well as lack of motivation, fuzzy orientation, and a low sense of identity, which could influence the inability of students to adapt to blended teaching [13].

Achievement is an important indicator to examine blended teaching effectiveness. In an educational process, students are considered to be successful if they can complete the education program on time with good learning outcomes. That is achievements are the realization of potential skills or capacity that students have. Learning achievements could be seen from students' mastery of the subjects they have taken. Studies on blended foreign language teaching have shown that blended teaching was effective in improving learners' foreign language skills [14]. Other findings also revealed that the impact of blended teaching on learners' effectiveness was positively predicted by achievement, engagement, involvement, retention, and cognitive outcome [15]. Lin and Gong found in their study that there was a significantly relationship between the initiative of university students to participate in blended teaching courses and students' expectation on their achievements in the course [16].

Additionally, some studies were carried out to evaluate the effects of gender on students' achievements in EFL blended teaching. Some researchers found that there was no significant difference in the achievements between male and female learners in blended teaching [17], some other researchers found that female students obtained significantly higher achievements than male students in EFL blended teaching environment [18]. The results need to be verified by data from the study on EFL blended teaching due to the findings of different blended teaching effectiveness among different disciplines.

Furthermore, according to the research findings of teaching conditions gap between urban and rural areas in China, the overall level of English teaching in rural areas is still relatively lower than that of the urban areas, and the gap is widening [19]. Students from rural areas have different starting points than their urban-based counterparts in English learning. This EFL educational gap between students from urban and rural areas will eventually affect students' further study and their development [20]. Other researchers found that there is a significant difference between college students from rural areas and college students from urban areas in their concept of mother tongue, and college students from rural areas have significantly higher level of concept of mother tongue than that among college students from urban areas [21]. Therefore, it is assumed that students from urban and rural areas would perform differently in EFL blended learning.

In order to solve the problems existing in the current EFL blended teaching in China, the research is conducted to examine the effects of an "SPOC and Flipped classroom"-based blended teaching strategy on the first-year English-major students' achievements in English language skills. "SPOC" was an abbreviation of "Small Private Online Course", which was originally proposed by Armando Fox in 2013 as a typical online teaching and learning paradigm in the "post-MOOC" era. In the expression, "Small" refers to the number of students generally ranging from dozens to hundreds. "Private", compared with "Open" in MOOC, means that only identified learners on campus have access to these resources and there exists a good level of course privacy [22]. SPOC places more emphasis on students having a complete and in-depth learning experience, which is conducive to improving the course completion rate. Teachers could set up and automatically regulate the teaching pace and course scoring management system of each course according to students' personal interests and preferences and actual needs. Flipped classroom teaching is a new teaching technology and method based on good network communication conditions in the 20th century. Originally, Lage et al. (2000) put forward the "Inverted Classroom" concept and defined it as "what traditionally happened inside the Classroom now happens outside the Classroom, and vice versa" [23] (p. 32). The concept involves students in more classroom activities, thus subverting the traditional classroom teaching schedule. The "Flipped Classroom" introduced by Baker (2016) emphasizes "understanding and application rather than recall; At the same time not sacrificing the presentation of factual knowledge; Students' having more control over their learning; Making students more responsible for their own learning; Providing more peer learning opportunities for students" [24] (p.9). Bishop and Verleger (2013) proposed the definition framework of flipped classroom and considered that flipped classroom consists of two parts, namely, interactive classroom activities guided by student-centered learning theory and the explicit instruction method guided by teacher-oriented learning theory [25]. This definition emphasizes the type of learning on which pre-class and in-class sessions are based and excludes classes that do not use video as learning material for extracurricular activities. Abeysekera and Dawson (2015) provided a general theoretical model for flipped classroom based on motivation theory and cognitive load theory. They believed that in flipped classroom, the feeling of competence, relevance, autonomy could improve learners' external and internal motivation, while customized expertise and self-paced learning can better manage learners' cognitive load [26].

1.1. Research Objectives and Research Questions

In the study, the "SPOC and Flipped-classroom"-based blended teaching strategy is implemented in the EFL blended teaching among first-year English-major students, and the effects of the blended teaching strategy on the first-year English-major students' achievements in blended teaching and learning environment are examined. The following research objectives were put forward: 1. To examine the first-year English-major students' achievements in the English language skills of speaking, reading, listening, translating, and writing using an "SPOC and Flipped classroom"-based blended teaching strategy. 2. To examine the first-year English-major students' achievements among male and female first-year English-major students using an "SPOC and Flipped classroom"-based blended teaching strategy. 3. To examine the first-year English-major students' achievements among first-year English-major students from urban and rural areas using an "SPOC and Flipped-classroom"-based blended teaching strategy.

According to the research objectives, the following three research questions (RQ) are put forward. 1: What are the first-year English-major students' achievements in English language skills using an "SPOC and Flipped classroom"-based blended teaching strategy? 2: What are students' achievements in English language skills among male and female first-year English-major students using an "SPOC and Flipped classroom"-based blended teaching strategy? 3: What are the first-year English-major students' achievements in English language skills among first-year English-major students from urban and rural areas using an "SPOC and Flipped classroom"-based blended teaching strategy?

1.2. Hypotheses of the Research

According to the research questions and the research objectives, the research hypotheses are as follows:

H1. *The "SPOC and Flipped classroom"-based blended teaching strategy has a significantly positive effect on the first-year English-major students' achievements compared with the traditional face-to-face classroom teaching strategy.*

H2. *The "SPOC and Flipped classroom"-based blended teaching strategy has a significantly positive effect on the first-year English-major students' achievements in listening compared with the traditional face-to-face classroom teaching strategy.*

H3. *The "SPOC and Flipped classroom"-based blended teaching strategy has a significantly positive effect on the first-year English-major students' achievements in reading compared with the traditional face-to-face classroom teaching strategy.*

H4. *The "SPOC and Flipped classroom"-based blended teaching strategy has a significantly positive effect on the first-year English-major students' achievements in translating compared with the traditional face-to-face classroom teaching strategy.*

H5. *The "SPOC and Flipped classroom"-based blended teaching strategy has a significantly positive effect on the first-year English-major students' achievements in writing compared with the traditional face-to-face classroom teaching strategy.*

H6. *The "SPOC and Flipped classroom"-based blended teaching strategy has a significantly positive effect on the first-year English-major students' achievements in speaking compared with the traditional face-to-face classroom teaching strategy.*

H7. *Male students have significantly higher achievements than those of female students using the "SPOC and Flipped classroom"-based blended teaching strategy.*

H8. *Students from urban areas have significantly higher achievements than those of students from rural areas using the "SPOC and Flipped classroom"-based blended teaching strategy.*

2. Materials and Methods

According to the research objectives, the quasi-experiment was conducted in the current research based on the philosophy of the quantitative methodology (See Figure 1).

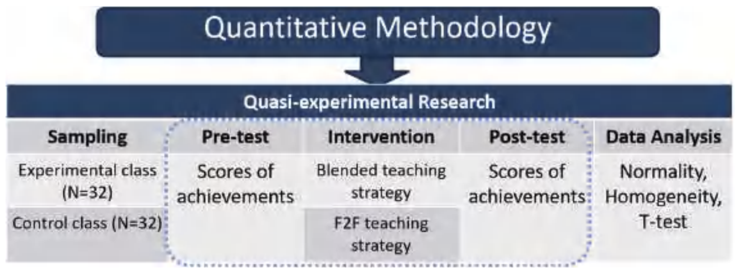


Figure 1. Research design.

2.1. Research Design of the Quasi-Experiment

A quasi-experiment was adopted in this research to examine the effects of "SPOC and Flipped classroom"-based blended teaching strategy to the first-year English-major students' achievements in five language skills, which consisted of a one-way two-group design comparing students' achievements between experimental and control classes. The rationale for using quasi-experiment in this study is that the essence of experimental design refers to the researcher's deliberate manipulation of or changes to the study situation of at least one independent variable in order to observe the effect of this on at least one response variable [27]. According to the research design of the quasi-experiment shown in Figure 1, the whole procedures of the quasi-experiment were composed of the following three sessions:

(1) The grouping of samples for the quasi-experiment: In order not to affect the normal teaching order, Class A (32 students) and Class B (32 students) of the first-year English-major students, who participate in the course of Integrated English (1) undertaken by the researcher in 2022, in the School of Foreign Languages of the university, are selected, respectively, as the control group and the experimental group. Class A (control group) and Class B (experimental group) have the same teaching content, teachers, and teaching hours (200 min/week).

(2) The conduction of the quasi-experiment (in the dotted box): The quasi-experimental research is conducted by implementing an intervention of "SPOC and Flipped classroom"-based blended teaching strategy in Class B (experimental group), while in Class A (control group), the traditional face-to-face teaching strategy was continuously applied. Initially, both in the experimental class and the control class, the researcher clarifies the teaching and learning goals of the course and those of each unit to students before starting the instruction. The purpose of this section is to let the students know clearly what they are going to do in learning each session. The goals of the course include knowledge aims, ability aims, and quality aims, among which the knowledge aims and ability aims of improving students' English language proficiency in five basic skills were the main objects of the research. Then, detailed explanation is given to students in line with the purpose, function, and significance of the experimental research, so as to enable students to be aware of blended teaching and make preparation for the smooth implementation of the quasi-experiment.

Before the quasi-experiment, a pre-test was conducted to obtain a general idea of the students' English knowledge and ability background in Class A (control group) and Class B (experimental group). The students' total achievements including the five English language skills are tested by questions in the pre-test paper, and the normal distribution of the data is tested by the descriptive analysis tool of SPSS and the results are reported, including minimum, maximum, mean, standard deviation, skewness, and kurtosis for illustrate the relevant information of the samples in the quasi-experiment to ensure the normality of the data and the reliability and validity of the inferential analysis results. Then, an independent sample *t*-test is conducted on the pre-test performance data to ensure that there is no significant difference between the initial level of students in the control group and the experimental group.

Next, the intervening teaching using the blended teaching strategy is conducted in the experimental class in the following stages. Firstly, "SPOC and Flipped classroom"-based blended teaching strategy is applied in the blended instruction to rearrange the teaching procedures for effective integration of online and face-to-face teaching and learning. The blended teaching and learning technologies provided by the "SPOC and Flipped classroom" approach are made full use of as the online and offline instruction are carried out. The teaching and learning session for each class include two instances (four periods) a week, amounting to 200 min in total. Then, the researcher provided the participants in the experimental class with an overview of the intervention in the course, including the teaching purpose, its procedures, and the general activities dealing with the course learning. A brief introduction of the term and model of "blended teaching", as well as the online learning platform including the features, functions, and usage, are provided for the learners in order to prepare them to successfully perform the activities that would be covered within the intervention sessions. The purpose of the introduction is to help the learners easily obtain access to the blended learning context by assuring them that the online platform is designed to be a user-friendly interface. Some concepts such as "SPOC" and "Flipped classroom" are introduced to the learners for their successful autonomous blended learning. Finally, the two different teaching and learning procedures based on the traditional face-to-face classroom teaching strategy and "SPOC and Flipped classroom" blended teaching strategy are carried out, respectively, in the control class and the experimental class.

After the intervention in the quasi-experiment, the post-test is carried out in the control class and experimental class. Participants' achievements in post-test are compared with those in pre-test to determine their respective differences.

(3) Statistical analysis on the data of tests in the quasi-experiment: Since the samples selected in this study are small and the scores are characterized by continuity, the statistical approaches could be used by a data analysis tool, Statistical Product and Service Solutions (SPSS) 26.0, to analyze the data and find out whether there were significant differences in students' English language achievements before and after the tests between the control class and the experimental class.

At the very beginning, the data are tested for homogeneity of variance to find out whether the variance difference between the two samples do not exceed the range specified by statistics between the two samples. If the variance of the two samples is not uniform, then the correction formula of mean comparison is used for processing [28].

Then, the normality of the data is also tested by the statistical method of frequency. According to literature, if the $|\text{skewness}| > 2$ or $|\text{kurtosis}| > 7$, then the data does not form a normal distribution [29,30].

Finally, *t*-tests are used on the data analysis for obtaining results to verify the research hypotheses and to answer the research questions. Before and after the intervention, independent sample *t*-tests are performed respectively on the scores of pre-test between control and experimental classes, as well as on those of post-test between the two classes to find out whether there are significant differences in students' achievements between the two classes. A paired-sample *t*-test is also adopted to test whether there are significant differences in students' achievements before and after the intervention, respectively, in the control class and the experimental class.

2.2. Blended Teaching Design Based on "SPOC and Flipped Classroom"-Blended Teaching Strategy

In the current research, the blended teaching procedures for the Integrated English (I) course is redesigned by implementing the "SPOC and Flipped-classroom"-based blended teaching strategy for deeply integrating the online and face-to-face instruction inside and outside of the classroom. The strategy is used to offer guidance, assistance, and resources both online and offline in blended teaching practice to the teacher and students through the techniques of SPOC and flipped classroom. The "SPOC and Flipped-classroom"-based blended teaching strategy is conducted during the teaching design in the following two aspects: To combine online and offline instruction by integrating SPOC with face-to-face classroom teaching in the flipped classroom teaching process for constructing a full time, all-round, fully guided, closed-loop learning environment for students. SPOC could enable teachers to provide abundant and appropriate learning materials, giving guidance, initiating interactions, and making evaluations of students with targeted online techniques. Flipped classroom could make up for the shortage of time and space in traditional classroom teaching, and at the same time can promote students' independent learning. On the other hand, this is to carry out instructional activities (both of teacher and students) with the scaffolding of "SPOC and Flipped classroom" techniques in the blended teaching procedures. The blended teaching design based on "SPOC and Flipped-classroom" blended teaching strategy is shown in Figure 2.

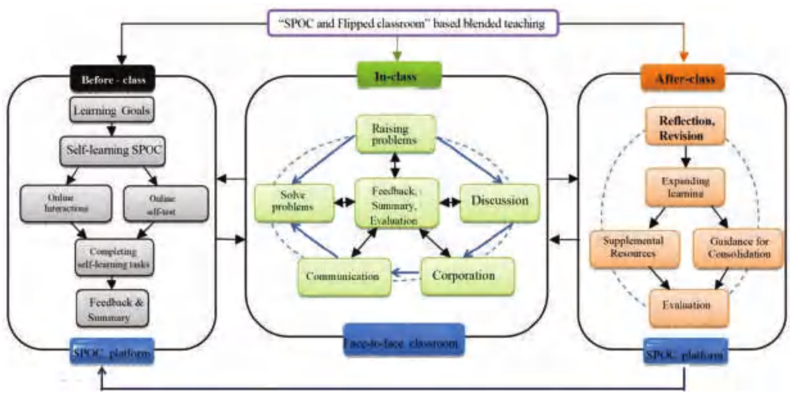


Figure 2. "SPOC and Flipped classroom"-based blended teaching design.

As is shown in Figure 2, the whole blended teaching process is divided into three sections—the before-class section, the in-class section, and the after-class section. Before the class, learning goals are released to students before the classroom teaching activities take place. This step is followed by students' self-learning based on SPOC, during which the interaction and self-test learning activities should be carried out by students under the teacher's guidance in the form of online tasks through the SPOC platform. Then, as students finish all the tasks, they receive feedback and a summary from their teacher. The SPOC techniques are performed through an online instructional platform named "Xue Xitong", which is a free integrated mobile teaching, mobile learning, mobile reading, and mobile social communication application developed by Beijing Century Superstar Information Technology Development Co., LTD. in 2016 in China. The functions of restricted class access, releasing videos, pictures, documents, teacher–students–peer interactions, monitoring learning process, and evaluating students' learning outcomes could help teachers in offering guidance and conducting monitoring for students' autonomous learning before and after traditional face-to-face classroom teaching. The present adaptation of SPOC is to create an autonomous learning environment for students, enabling improvements in their learning process.

In-class activities are carried out by both the teacher and the students. The problems raised at the beginning of the class would be solved through discussion, cooperation, and communication in teacher–student and student–student interactions. Each link of the loop (see "in-class" section in Figure 2) can be adjusted at any time based on feedback, summary, and evaluation from both teachers and students. This section is performed as face-to-face classroom teaching, and is composed of the "discussion", "cooperation", "communication", and "problem-solving" steps after students' autonomous SPOC learning of the course in the flipped classroom, which is the reverse of the traditional face-to-face classroom teaching procedures.

After-class teaching and learning happen as soon as the in-class teaching and learning aims are accomplished. This section is conducted as the reflection and revision of in-class teaching and learning process. According to the evaluation of students' performance in knowledge acquisition and skill management with the aids of online and offline teaching techniques, a teacher would provide more personalized self-learning supplemental learning resources and guidance to students for the consolidation of the knowledge and skills acquisition. Then, students' performance in this section would also be evaluated. The section would be carried out on the SPOC platform as well.

The three sections would not stop as the after-class section finishes, but would repeat to achieve the blended teaching and learning goals of the course.

Meanwhile, the traditional face-to-face classroom teaching strategy is implemented in the teaching procedures of the control class, which is also designed as three teaching stages: before-class, in-class and after-class. See Figure 3.

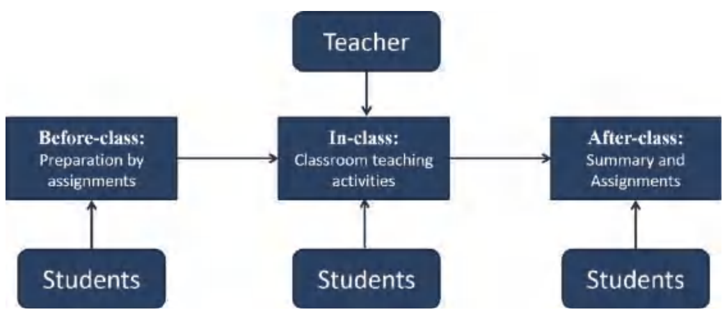


Figure 3. Traditional face-to-face classroom teaching design.

As Figure 3 shows, the traditional face-to-face classroom teaching approach was also carried out in three stages. The before-class stage refers to students' preparation for the subject of the new class (the "preview") according to the assignments given by the teacher before the teaching happens. The in-class stage includes all classroom teaching activities such as situation creating, knowledge interpreting, exemplifying, collaborating, and so on. The last stage involves the summary and assignments students had to work on after the class for reviewing the old knowledge and prepare for the new lesson (known as "review and preview"). The teacher only participates in the interactions with students in the in-class periods, and students have to complete the before-class and after-class tasks autonomously with the guides for assignments given by the teacher.

2.3. Population and Samples

The research design section follows the type of design with characteristics of the population and the sampling procedure. Within this target population, a sample was selected that consisted of first-year English-major students at an undergraduate university in Shandong Province of China (see Figure 4).

Figure 4. Population and samples.

The research sample was selected from the population using randomization and stratified sampling techniques according to probability sampling techniques. The samples who voluntarily agreed to take part in the research include both men and women who enrolled for the 2022 English-major courses in H University. Among the English-major courses, Integrated English (I) is the compulsory core course which begins in the first semester for first-year English-major students. These students participate in the course to improve their proficiency in the five basic English language skills — listening, reading, speaking, writing and translating. In the quasi-experiment, there were 64 English-major students from the same Chinese university in Shandong province, who were divided into the control class (N = 32, (intact group), with 6 men and 26 women) and the experimental class (N = 32, (intact group), with 4 men and 28 women).

2.4. Data Collection Instruments

The fundamental purpose of experimental design is to impose control over conditions that would otherwise cloud the true effects of the independent variables upon the dependent variables [6]. The quasi-experiment, in this research, is adopted to examine students' achievements by implementing both the "SPOC and "Flipped classroom" blended teaching strategy in the experimental class for the intervention, and the traditional face-to-face blended teaching strategy for the control class.

Testing is used as the instrument to collect the quantitative data of English-major students' achievements of English language competence — speaking, reading, listening, writing, and translating — before and after the blended teaching intervention in the course of Integrated English (I) for a semester (four months). The tests include a pre-test and a post-test.

2.4.1. Composition of the Test Papers

Each of the two test papers is composed of five types of questions—listening, reading, speaking, writing, and translating. The questions involving the five different language skills are cited from two large-scale, high-stake tests for measuring Chinese students' English language competence—English Test for International Communication (ETIC) and College English Test Band 4 (CET4).

The five types of questions compose the pre-test and post-test papers to measure the first-year English-major students' five basic English language skills.

(1) Listening

The listening section consists of five questions in the form of sentence completion. The listening materials is a conversation between two people in a context of social communication, which plays back for two times. The assessment of the listening test is based on the rating standard. Each correct answer receives 2 marks and scores are out of 10.

(2) Reading

The reading section is composed of two passages, and each passage includes five multiple-choice questions. The assessment of the reading test is based on the rating standard. Each correct answer receives 2 marks and scores are out of 20.

(3) Translating

The translating section consists of one task, which required participants to translate a short essay from Chinese to English. The assessment of translating test is based on accuracy, literal fluency and coherence in the lexical, syntactical and grammatical ranges, and the scores are out of 30.

(4) Speaking

The speaking section consists of two parts in the form of self-presentation. Students obtained the requirements of the speaking test when they received the test papers to obtain a general idea of the contents and requirements. In the first part (self-introduction), students are required to briefly introduce themselves freely within 2 min. In the second part (individual long turn), students talk about a particular topic which allows them to present any points they wish to cover. The assessment of the speaking test is based on fluency and coherence, lexical resource, grammatical range and accuracy, pronunciation, and intonation, and the scores are out of 30.

(5) Writing

The writing section consists of one task, which requires learners to write 120 words. In the task of the writing test, a situation is presented to the participants, and they are asked to write an essay providing information meeting the demands of the situation. The assessment of the writing test is based on task achievement, coherence and cohesion, lexical resource, and grammatical range and accuracy, and the scores are out of 20.

The pre-test and post-test were administered for the participants of both the experimental and the control group. The test papers were used in the pre-test and post-test to measure the participants' language proficiency of the five skills (listening, reading, speaking, writing and translating) before and after the intervention sessions. Each of the five raters scored one certain type of question (each of the five skills) included in the test papers independently according to the corresponding rating scale.

2.4.2. Reliability of the Test Papers

In order to test the reliability of the test, the researcher administered a pilot test on a sample which consisted of 30 first-year English-major students except for the samples, and the reliability was determined through Cronbach Alpha test. There were two tests (pre-test and post-test) and each of the tests took 110 min. The consistency coefficient was calculated using Cronbach Alpha and the result (pre-test: Cronbach's Alpha = 0.702; post-test: Cronbach's Alpha = 0.742). Thus, reliability of the test was deemed acceptable.

2.5. Data Analysis

The quantitative analysis on the data was conducted through SPSS 26.0. The procedure of data analysis contained the following steps.

Firstly, normality was tested to determine the normal distribution of the data. Students' total achievements including the five English language skills were tested in the post-test, and the normal distribution of the data was tested by the descriptive analysis tool of SPSS and the results are reported too, including minimum, maximum, mean, standard deviation, skewness and kurtosis for illustrate the relevant information of the samples in the quasi-experiment to ensure the normality of the data and the reliability and validity of the inferential analysis results of the post-test data. Then, an independent sample *t*-test was conducted on the post-test performance data to find out whether there was significant difference between in students' achievements both in the control group and the experimental group, and before and after the intervention.

Then, homogeneity of variance of the data was tested to ensure that the variance of two samples was homogeneous.

Finally, in order to ensure that there is no significant difference between the experimental class and control class regarding their language learning skills at the beginning of the study, an independent sample *t*-testwas performed.

Specifically, the data of the quasi-experiment were collected through pre-test and post-test and then were analyzed by performing independent sample *t*-test and paired-sample *t*-test in SPSS 26.0. The results, such as the maximum, minimum, mean, *t*-value, Sig. (2-tailed), mean difference, and *p*-value, were reported to verify whether there was a significant difference between the effects of the blended learning strategy and those of the traditional face-to-face teaching strategy on students' achievements. The independent sample *t*-test was used for verifying whether there was any difference in students' achievements between the control class and the experimental class, between male and female students, and between students from urban areas and students from rural areas. ANCOVA (analysis of covariance) was conducted for eliminating the influence of confounding factors (covariates of pre-test and the two teaching strategies) on the analysis index and identifying whether there was any statistically significant correlation between the independent and dependent variables following the intervention.

3. Results

In this section, the results of exhaustive data analyses are presented. These were performed to answer the research questions and to verify the research hypotheses, and the results of the statistical analysis were provided.

3.1. Descriptive Statistics of the Quasi-Experiment

The demographic information of the samples in the experimental class, including numbers, genders and regional backgrounds, was analyzed in descriptive statistics. The distribution of genders and regional backgrounds among sample of participants is shown in the following Table 1.

Table 1. Demographic information of the samples.

			Gender		
		Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Valid	1	6	18.8	18.8	18.8
	2	26	81.2	81.3	100.0
	Total	32	100.0	100.0	
			Regional Background		
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	7	21.9	21.9	21.9
	2	25	78.1	78.1	100.0
	Total	32	100.0	100.0	

Then, normal distribution tests of the statistics in the pre-test and post-test were carried out for the control class and the experimental class of the quasi-experiment, and the results are shown in Tables 2 and 3.

Table 2. Descriptive statistics of students' achievements in the pre-test.

	Valid N	Minimum	Maximum	Mean	Std. Deviation	Skewness	Kurtosis
EC	32	51	81	67.47	6.258	0.069	0.762
CC	32	56	81	68.19	6.631	0.293	−0.556

EC = experimental class, CC = control class.

Table 3. Descriptive statistics of students' achievements in the post-test.

	Valid N	Minimum	Maximum	Mean	Std. Deviation	Skewness	Kurtosis
EC	32	61	87	75.72	5.887	−0.432	0.286
CC	32	41	89	65.41	10.121	−0.044	0.263

EC = experimental class, CC = control class.

As is shown in Table 2, the EC had a slightly lower mean (67.47) and a smaller standard deviation (6.258) compared to CC, which has a mean of 68.19 and a larger standard deviation of 6.631. As for the skewness and kurtosis, the skewness of the experimental class is 0.069. A positive value indicates a right-skewed distribution, and a negative value indicates a left-skewed distribution. In this case, the value is close to zero, suggesting a nearly symmetrical distribution. Kurtosis of the experimental class is 0.762 a positive value indicates heavier tails compared to a normal distribution, and a negative value indicates lighter tails. A value close to zero suggests a distribution similar to a normal distribution. The positive value of skewness of the control class is 0.293, indicating a slight right-skewed distribution, but skewness is not substantial. The kurtosis of the control class is −0.556. The negative value indicates that the distribution has lighter tails compared to a normal distribution. $|\text{skewness}| < 2$ and $|\text{kurtosis}| < 7$ were found in both experimental and control classes, which means the data of pre-test the control class form a normal distribution.

As can be seen from Table 3, the scores of the experimental class had a higher mean (75.72) and a smaller standard deviation (5.887) compared with the scores of the control class, which had a mean of 65.41 and a larger standard deviation of 10.121. According to the statistics, the skewness of experimental class is −0.432. The negative value indicates a left-skewed distribution, meaning the tail is extended towards the left side of the distribution. The kurtosis of the experimental class is 0.286. The positive value indicates heavier tails compared to a normal distribution. The value is close to zero, suggesting a distribution similar to a normal distribution. The skewness of control class is −0.044, which is close to zero, indicating a nearly symmetrical distribution for control class. The kurtosis of control class is 0.263. The value is close to zero, suggesting a distribution similar to a normal distribution. $|\text{skewness}|$ is less than 2 and $|\text{kurtosis}|$ is less than 7 for the statistics of both experimental class and control class in the post-test, which means the data of post-test in the experimental class and control class form a normal distribution.

3.2. Quantitative Analysis Results and Findings

The quantitative analysis of students' achievements was conducted based on the quantitative data collected by tests. Scores of both the pre-test and the post-test in the controlled class and experimental class were analyzed by SPSS 26.0. Different statistical methods were used in the two phases to obtain the results, with the aim of elucidating the research questions and verifying the hypotheses.

The quasi-experiment was designed to find out the effect of the "SPOC and Flipped classroom"-based blended teaching strategy on students' achievements through a comparison with the effect of the face-to-face classroom teaching strategy; therefore, it should elucidate whether there was a significant difference between the independent variables ("SPOC and Flipped classroom"-based blended teaching strategy vs. traditional face-to face classroom teaching strategy) and the dependent variable (students' achievements).

In this case, the statistical analysis methods such as the independent sample *t*-test and the paired sample *t*-test were conducted on students' scores in pre- and post-tests to examine the effects of the two blended teaching strategies on students' achievements. In addition, the moderating effects of gender and regional background on students' achievements also were examined under the conditions of different teaching strategies being applied in EFL instruction.

3.2.1. Results and Findings for Verifying H1, H2, H3, H4, H5, and H6 involving Research Question 1

The results and findings from the quantitative data analysis of the quasi-experiment were derived in order to answer the first research question and to verify the research hypotheses 1–6.

The dependent variables involved in the analysis were students' achievements in five English language skills, including listening, reading, translating, writing and speaking. The independent variables were two different teaching strategies —the "SPOC and Flipped classroom"-based blended teaching strategy and the face-to-face classroom teaching strategy.

In order to ensure that there was no significant difference between the control group and the experimental group regarding their language learning skills at the beginning of the research, the pre-test was conducted synchronously in both experimental class and control class, and then an independent sample *t*-test was performed on the scores. The results are provided in Table 4.

Table 4. Results from independent sample *t*-test on pre-test scores.

	EC (n = 32)		CC (n = 32)				
Pre-test	M	SD.	M	SD	Sig. (2-tailed)	MD	t
	67.47	6.258	68.19	6.631	0.657	−0.72	−0.446

Through the independent sample *t*-test on the scores collected from the pre-test in the two classes, it was found that they were not any significant differences among the students of the both classes in terms of their achievements in five language skills (listening, reading, translating, writing and speaking) (t = −0.446, $p > 0.05$). Inspections of the means indicate that the average pre-test scores of students in the two classes were at an almost at the similar level.

After the intervention sessions, the students in both classes took part in a post-test. The major purpose of this step was figuring out that there was significant difference between the two classes after the intervention in the experimental class; therefore, the independent sample *t*-test was performed to analyze the scores collected from the two classes from the post-test. The statistics including a valid number of samples, mean, standard deviation, and the inferential statistics, are provided in Table 5.

Table 5. Results from independent sample *t*-test on post-test scores.

	EC (n = 32)		CC (n = 32)				
Post-test	M	SD	M	SD	Sig. (2-tailed)	MD	t
	75.72	5.887	65.41	10.121	0.000 *	10.31	4.982

* $p < 0.05$.

The results in Table 5 show that the "SPOC and Flipped classroom"-based blended teaching strategy had a significantly positive effect on the participating first-year English-major students' achievements compared with the traditional face-to-face classroom teaching strategy. Inspections of the two classes means indicated that the average post-test score of students learning in blended teaching designed based on the "SPOC and Flipped classroom"-based blended teaching strategy is significantly higher than the score of students learning in face-to-face classroom teaching designed on the basis of the face-to-face classroom teaching strategy (EC(M) = 75.72, CC(M) = 65.41, $p < 0.05$). From the results, Hypothesis 1 is accepted. The result is consistent with the findings of Bilgin [6] and Oweis [31]: blended teaching could be helpful in improving students' English language skills.

The effective size of the pre- and post- test in both of the the control class and the experimental class was tested using the Cohen's d. See Table 6.

Table 6. Independent sample effect size.

		Normalized Quantitya	Point Estimation	95% Confidence Interval	
				Minimum	Maximum
pre-test CC-EC	Cohen d	6.447	0.111	−0.379	0.601
	Hedges revised	6.526	0.110	−0.375	0.594
	Glass Delta	6.258	0.115	−0.377	0.605
post-test CC-EC	Cohen d	8.280	−1.246	−1.778	−0.705
	Hedges revised	8.381	−1.230	−1.756	−0.696
	Glass Delta	5.887	−1.752	−2.398	−1.089

Denominator used when estimating the size of the effect. Cohen d convergence standard deviation. Hedges: The correction uses the convergence standard deviation plus the correction factor. Glass Delta: The sample standard deviation of the control group was used.

The results in in Table 6 show that the effective size of the pre-test in both the control and the experimental classes, as measured by Cohen's d shown, is d = 0.111, indicating a low effect. The effective size of the post-tests among the two classes, as measured by Cohen's d, is d = 1.246, indicating a high effect.

Furthermore, in order to determine whether there is significant difference in students' achievements before and after the intervention session within the same class (in the experimental class and in the control class, respectively), two paired-sample *t*-tests were conducted on the pre-test and post-test scores collected from the both the classes. The results can be seen in Table 7.

Table 7. Paired-sample *t*-test of control and experimental classes.

	Pre-test of CC (n = 32)		Post-test of CC (n = 32)				
Students' Achievements	M	SD	M	SD	Sig. (2-tailed)	MD	t
	68.19	6.631	65.41	10.121	0.117	2.78	1.611
	Pre-test of EC (n = 32)		Post-test of EC (n = 32)				
	M	SD.	M	SD	Sig. (2-tailed)	MD	t
	67.47	6.258	75.72	5.887	0.000 *	−8.25	−6.955

* $p < 0.05$.

The results in Table 7 show that there is no significant difference in students' achievements between pre-test and post-test in the control class (t = 1.611, $p > 0.05$). Inspections of the two tests' means indicated that the average score of students' achievements had no significant improvement before and after the intervention of face-to-face classroom teaching strategy. The results show that students' achievements in the pre-test is significantly different from the students' achievements in the post-test in the experimental class ($p < 0.05$). Inspections of the two tests' means indicate that the average score of students' achievements in the pre-test was significantly lower than that in the post-test. This means that students' achievements was improved after the intervention with the "SPOC and Flipped classroom"-based blended teaching strategy. The results shown in Tables 4 and 5 verify H1: the "SPOC and Flipped classroom"-based blended teaching strategy has a significantly positive effect on the first-year English-major students' achievements. This finding is consistent with that of Bañados' (2006) study on the improvement of Chilean undergraduate students' EFL language skills in a blended teaching environment [32].

In order to eliminate the influence of confounding factors (covariates of pre-test) on the analysis index and identify whether there were any statistically significant correlations between the independent and dependent variables following the intervention, results from the pre-test and post-test of students' achievements in the experimental class were also evaluated by a mixed-design ANCOVA (analysis of covariance). The results are shown in Table 8.

Table 8. ANCOVA on data of experimental class.

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	2728.797 a	3	909.599	16.934	0.000
Intercept	1226.839	1	1226.839	22.839	0.000
Teaching Strategy "1blend2facetoface"	1633.151	1	1633.151	30.404	0.000 *
Pretest	457.726	1	457.726	8.521	0.005
Error	3222.953	60	53.716		
Total	324,612.000	64			
Corrected Total	5951.750	63			

* $p < 0.05$. a: R Squared = 0.458 (Adjusted R Squared = 0.431).

The analysis of covariance in Table 8 shows that the significance of the "Teaching Strategy" variable was less than 0.05. This suggested the presence of a significant difference ($\alpha = 0.05$) among the means of the achievement post-test results in the items of the overall achievement test in the five English language skills; these differences can be attributed to the teaching strategy variable ("SPOC and Flipped classroom"-based blended teaching strategy versus traditional face-to-face teaching strategy). This indicates the presence of an effect for the blended teaching strategy on the achievement of first-year English-major students' five English language skills.

In order to conduct an in-depth exploration of the findings about the effects of the blended teaching strategy on students' achievements in different English language skills to verify H2, H3, H4, H5, and H6, further statistical analysis on the data was performed. The descriptive statistics and inferential statistics of the independent sample *t*-test were conducted on the scores of the listening, reading, translating, writing, and speaking pre-test and post-test for the experimental and control classes. The results are provided in Tables 9 and 10.

The purpose of the independent sample *t*-test of the the pre-test results was to ensure that there was no significant difference in students' achievements of the five English language skills between the control and experimental classes before the intervention of teaching strategies. The results of Levene's test of the pre-test scores of each skill among the control and experimental classes demonstrated that the variables satisfied the homogeneity of variance.

As is shown in Table 9, there are no significant differences in students' achievements of five English language skills between the control class and the experimental class in the pre-test (listening: $t = -1.60$, $p > 0.05$; reading: $t = 0.170$, $p > 0.05$; translating: $t = 1.179$, $p > 0.05$; writing: $t = 1.434$, $p > 0.05$; speaking: $t = 0.242$, $p > 0.05$). This means that before the interventions with the two strategies, students' English language skills in listening, reading, translating, writing and speaking were at a similar level in both the control class and the experimental class.

Table 9. Independent sample *t*-test on pre-test results.

	Groups	N	Mean	SD	Sig. (2-Tailed)	MD	t
Listening	CC	32	5.81	2.235	0.113	−0.94	−1.609
	EC	32	6.75	2.423			
Reading	CC	32	12.88	2.959	0.865	0.13	0.170
	EC	32	12.75	2.907			
Translating	CC	32	13.75	2.410	0.243	0.72	1.179
	EC	32	13.03	2.469			
Writing	CC	32	14.94	2.047	0.156	0.69	1.434
	EC	32	14.25	1.778			
Speaking	CC	32	20.81	2.494	0.809	0.12	0.242
	EC	32	20.69	1.512			

Table 10. Independent sample *t*-test on post-test results.

	Groups	N	Mean	SD	Sig. (2-Tailed)	MD	t
Listening	CC	32	4.25	2.627	0.010 *	−1.63	−2.648
	EC	32	5.88	2.268			
Reading	CC	32	12.38	3.883	0.007 *	−2.43	−2.780
	EC	32	14.81	3.084			
Translating	CC	32	13.63	3.998	0.000 *	−3.06	−3.854
	EC	32	16.69	2.055			
Writing	CC	32	12.66	3.288	0.000 *	−3.12	−4.905
	EC	32	15.78	1.475			
Speaking	CC	32	22.50	2.258	0.912	−0.06	−0.111
	EC	32	22.56	2.257			

* $p < 0.05$.

The results in Table 10 show that the "SPOC and Flipped classroom"-based blended teaching strategy has a significantly positive effect on the first-year English-major students' achievements in listening compared with the traditional face-to-face classroom teaching strategy, (EC(M) = 5.88, CC(M) = 4.25, $p < 0.05$). From this finding, Hypothesis 2 is accepted, and this is consistent with the findings of Aji [33] in the research.

As for reading skills, the results in Table 10 show that the "SPOC and Flipped classroom"-based blended teaching strategy had a significantly positive effect on the first-year English-major students' achievements in reading compared with the traditional face-to-face classroom teaching strategy ($p < 0.05$, EC(M) = 14.81, CC(M) = 12.38). From the finding, Hypothesis 3 was accepted and this is consistent with the findings of Bolandifar [34], Ghazizadeh and Fatemipour [35] and Yudhana [36].

As can be seen from Table 10, statistics shows that the "SPOC and Flipped classroom"-based blended teaching strategy had a significantly positive effect on the first-year English-major students' achievements in translating compared with the traditional face-to-face classroom teaching strategy ($p < 0.05$, EC(M) = 16.69, CC(M) = 13.63). From this finding, Hypothesis 4 was accepted.

Results in Table 10 show that "SPOC and Flipped classroom"-based blended teaching strategy has a significantly positive effect on the first-year English-major students' achievements in writing compared with the traditional face-to-face classroom teaching strategy ($p < 0.05$, EC(M) = 15.78, CC(M) = 12.66). From this finding, Hypothesis 5 was accepted, and the finding is also corroborated the findings of Hamouda that students of the blended learning group significantly outperformed the control group in their writing performance [37].

As is shown in Table 10, the results show that the "SPOC and Flipped classroom"-based blended teaching strategy had no significantly positive effect on the first-year English-major students' achievements in speaking compared with the traditional face-to-face classroom teaching strategy ($p > 0.05$). From this finding, Hypothesis 6 was rejected. This finding is opposite to the findings of Ginaya et al. [38], who found that the students participating in a blended teaching and learning model significantly improved in their speaking ability compared with those participating in a conventional face-to-face classroom teaching model.

3.2.2. Results and Findings for Verifying H7 Involving Research Question 2

The independent variables of the research were the two different teaching strategies—"SPOC and Flipped classroom"-based blended teaching strategy and the traditional face-to-face classroom teaching strategy. The dependent variable to be tested in the first research question was students' overall achievements in five English language skills. In addition, gender and regional background were involved in the research as moderating variables. In order to verify the seventh hypothesis, an independent sample *t*-test was performed twice to identify the effects of the moderating variable of genders on students' achievements. The statistics of both classes from the pre-test and post-test, in relation to gender, are shown in Table 11.

Table 11. Students' Achievements between Genders.

	M of EC (n = 6)		F of EC (n = 26)				
Students' Achievements in Pre-test (EC)	M	SD.	M	SD	Sig. (2-tailed)	MD	t
	68.83	6.616	67.50	6.320	0.647	1.333	0.462
	M of EC (n = 6)		F of EC (n = 26)				
Students' Achievements in Post-test (EC)	M	SD.	M	SD	Sig. (2-tailed)	MD	t
	78.67	7.737	75.04	5.333	0.178	3.628	1.380

M = male students; F = female students.

The independent sample *t*-test was performed to ensure that the results after the intervention would not be a product of the differences between the mean scores of the classes before conducting the experiment. As was shown in Table 11, the results show that male students did not have significantly higher achievements than female students before and after participating in the "SPOC and Flipped classroom"-based blended teaching strategy ($p > 0.05$). From this finding, Hypothesis 7 was rejected, and it is consistent with the finding of BENHADJ (2021) [39], that no significant gender differences in students' achievements were observed in blended teaching. Moreover, due to the difference in the number of male and female students, the effective size of the independent sample was tested. Results in in Table 12 show that the effective size of the independent sample, as measured by Cohen's d, is d = 0.625, indicating a medium effect [40].

3.2.3. Results and Findings for Verifying H8 Involving Research Question 3

Regional background was also considered in the research as a moderating variable to evaluate students' achievements before and after intervention through the "SPOC and Flipped classroom"-based blended teaching strategy in the experimental class for verifying the eighth hypothesis.

The scores of the pre-test and the post-test were analyzed using an independent sample *t*-test through SPSS 26.0, and the results are shown in Table 13.

Table 12. Independent sample effect size.

		Normalized Quantity	Point Estimation	95% Confidence Interval	
				Minimum	Maximum
Pre-test	Cohen d	6.361	−0.026	−0.914	0.862
	Hedges revised	6.526	−0.026	−0.891	0.840
	Glass Delta	6.320	−0.026	−0.914	0.862
Post-test	Cohen d	5.803	0.625	−0.281	1.522
	Hedges revised	5.953	0.609	−0.274	1.483
	Glass Delta	5.333	0.680	−0.233	1.581

Denominator used when estimating the size of the effect. Cohen d convergence standard deviation. Hedges: The correction uses the convergence standard deviation plus the correction factor. Glass Delta: The sample standard deviation of the control group was used.

Table 13. Students' Achievements between Regional Backgrounds.

	Urban (n = 7)		Rural (n = 25)				
Students' Achievements in Pre-test (EC)	M	SD.	M	SD	Sig. (2-tailed)	MD	t
	67.71	5.376	67.40	6.583	0.909	0.31	0.314
	Urban (n = 7)		Rural (n = 25)				
Students' Achievements in Post-test (EC)	M	SD.	M	SD	Sig. (2-tailed)	MD	t
	79.71	4.386	74.60	5.831	0.040 *	5.114	2.146

* $p < 0.05$.

This independent sample *t*-test conducted on the pre-test is to ensure that the results before the intervention would not be attributed to the differences between the mean scores of the classes before conducting the experiment. From the statistics shown in Table 13, the scores of students from urban areas showed no significant differences when compared with those of the students from rural areas in the pre-test ($p > 0.05$). The results of post-test scores show that students from urban areas have significantly higher achievements than those of students from rural areas participating in the "SPOC and Flipped classroom"-based blended teaching strategy ($p < 0.05$, urban (M) = 79.71, rural (M) = 74.60). From this finding, Hypothesis 8 is accepted.

Due to the difference in the number of students from urban areas and students from rural areas, the effective sizes of the independent sample data in the pre-test and the post-test were tested. The results are shown in Table 14.

Table 14. Independent sample effect size.

		Normalized Quantity	Point Estimation	95% Confidence Interval	
				Minimum	Maximum
Pre-test	Cohen d	6.360	0.049	−0.789	0.887
	Hedges revised	6.525	0.048	−0.769	0.865
	Glass Delta	6.583	0.048	−0.791	0.885
Post-test	Cohen d	5.572	0.918	0.041	1.780
	Hedges revised	5.716	0.895	0.040	1.735
	Glass Delta	5.831	0.877	−0.005	1.742

Denominator used when estimating the size of the effect. Cohen d convergence standard deviation. Hedges: The correction uses the convergence standard deviation plus the correction factor. Glass Delta: The sample standard deviation of the control group was used.

As is seen in Table 14, the effective size of the independent sample, as measured by Cohen's d, is d = 0.918, indicating a high effect [40].

4. Discussion

The results of the quantitative data emphasized the effects of the "SPOC and Flipped classroom"-based blended teaching strategy implemented in EFL instruction on the achievements of first-year English-major students. The following major findings were derived.

4.1. The "SPOC and Flipped Classroom"-Based Blended Teaching Strategy Was Effective in Improving Students' Overall Achievements

The effects of the "SPOC and Flipped classroom"-based blended teaching strategy on the first-year English-major students' achievements were tested through a comparison with the effects among students participating in a traditional face-to-face classroom teaching strategy. The results of students' overall achievements in English language skills indicate that the adoption of the "SPOC and Flipped classroom"-based blended teaching strategy as scaffolding led to an increase in English language achievements among the first-year English-major learners compared with those using the traditional face-to-face classroom teaching strategy. The result implies that the "SPOC and Flipped classroom"-based blended teaching strategy is more advantageous in helping Chinese EFL learners acquire both English language knowledge and improve English language skills in comparison with traditional face-to-face classroom teaching.

In order to eliminate the pretest differences in performance and identify the statistical significance of the difference according to the variable of teaching strategies, the analysis of covariance was used. The analysis of covariance suggested that the presence of a significant difference between the means from the achievement post-test in the items of the achievement test in English language skills was attributable to the teaching strategy variable (blended teaching versus face-to-face). This indicates the presence of an impact for the blended teaching strategy on the achievement of the first-year English-major students.

Therefore, when English language instruction was decoded using blended strategies as the scaffolding in EFL teaching and learning, language learning became easier for learners. The success of the "SPOC and Flipped classroom"-based blended teaching strategy was due to two remarkable reasons: firstly, its effectiveness in conducting the leaning process of EFL; its role in creating a different and more effective blended learning environment for students' learning context, collaboration, conversation, and meaning construction in comparison with the traditional face-to-face classroom teaching strategy, both for the learners and the teachers. The "SPOC and Flipped classroom" approach was used as scaffolding in the blended teaching of Integrated English (I), which created an effective learning environment and enabled the students to better access skills of meaning construction. Studies have shown that a good learning environment can provide learners with a variety of learning styles, rich course resources, convenient communication channels, and timely feedback and evaluation; these factors can motivate learners to be more actively involved in the course, and further promote their learning achievements and satisfaction [41].

4.2. "SPOC and Flipped Classroom"-Based Blended Teaching Strategy Has Significant Positive Effects on First-Year English-Major Students' English Language Skills of Listening, Reading, Translating, and Writing

The findings of the further data analysis on achievements in the five English language skills have different implications for the current blended teaching approach.

The significant difference observed from the statistical analysis results of listening, reading, translating and writing between the experimental and control class show that the first-year English-major students have significantly better performance in the English language abilities of listening, reading, translating, and writing from participating in the "SPOC and Flipped-classroom"-based blended teaching strategy in comparison with the traditional face-to-face classroom teaching strategy. From the significant results of higher scores of the four language skills in the experimental class, we can derive that the "SPOC and Flipped classroom"-based blended teaching strategy, used as the scaffolding in blended teaching, had significantly positive effects in improving students' English language skills of listening, reading, translating, and writing.

Alnoori and Obaid (2017) note that blended learning appears to be more effective than traditional methods because it has the flexibility to combine a range of techniques. In addition, the technology used in blended learning environments can be promoted and moderated by teachers in the classroom to prevent the technology from being misused or used in ineffective ways [42]. Based on the "SPOC and Flipped classroom"-based blended teaching strategy, as soon as the learning aims and tasks were assigned to students before each class, support becomes available for students in their online autonomous learning, through online learning materials such as audio and video materials provided by the teacher. Then, the in-class activities were carried out as problem-solving and evaluating stages in the face-to-face classroom teaching periods. Additionally, this strategy enables cooperation between the teacher and the students, as well as among students themselves. Finally, the evaluation of students' outcomes and performance of each listening practice is delivered in the form of evaluation and feedback after face-to-face class interactions. In this way, students' subjective initiative can adjust for exploratory study and learning, and they build an overall grasp of what they have learned.

4.3. The "SPOC and Flipped Classroom"-Based Blended Teaching Strategy Has No Significantly Positive Effects on First-Year English-Major Students' Achievements in EFL Speaking

According to the results of students' achievements in the speaking skill, no significant difference was found between the experimental class and the control class ($p > 0.05$). This finding indicated that the "SPOC and Flipped classroom"-based blended teaching strategy had no significant effect on the first-year English-major students' achievements of speaking in EFL blended learning.

The finding in this research was not consistent with the findings of some previous studies. Kirgoz (2011) found that there was a significant improvement in oral communication skills of the student teachers of English after a blended strategy and they had a positive perception of the application of the blended learning [43]. Similar findings have been reported by Yang et al. (2013) [44], Hung (2015) [45], and Alshumaimeri and Almasri (2012) [46], who found that blended learning had a significant effect on students' English listening and speaking, learning outcomes, and reading comprehension. The findings of Ginaya et al.'s research (2018) revealed that the students participating in the treatment of blended teaching significantly improved in their English-speaking abilities, and the improvement was also supported by their increased learning motivation and interest [37]. There are various reasons that can lead to different developments in students' EFL speaking abilities. The finding of the present study is contrary to the findings of Ginaya et al. (2018) [37], who found that the students participating in the blended teaching strategy were significantly improved in terms of their English-speaking ability compared with those participating in the conventional teaching model. Tang (2005) [47] suggested that the instrumental motivation and duration of oral English studies could influence students' speaking performance, and that the method of phonetic and oral teaching had an important influence on the motivation of oral learning. Wang (2007) also referred to students' learning motivation as one of the main factors affecting the improvement of oral English ability [48]. In addition, the language learning environment was found to be an influencing factor that affected students' speaking learning, because the traditional "teacher-centered" teaching environment affected and inhibited students' active participation in speaking learning activities. Zhang (2017) [49] used an oral English teaching practice that utilized blended learning—combining a web-based teaching method with traditional face-to-face teaching modes, integrated a variety of teaching equipment, and made language teaching more convenient. In addition to the possible reasons mentioned in the previous studies, the different results that were generated in the present study could be explained by the following factors: One factor is the different samples in studies. The samples in the previous studies were non-English-major students, while in this research they were first-year English-major students. The differences in the duration of English-speaking learning, and the environment between English-major and non-English major students, could lead to their different levels of speaking improvements. The other factor is the different strategies used in English-speaking blended teaching. Although researchers in China and other countries have conducted numerous studies on the effects of blended teaching in EFL instruction, "blended teaching" is still different, utilizing different strategies, which could cause different results in different studies.

4.4. Gender Has No Significantly Moderating Effect on the First-Year English-Major Students' Achievements in Their EFL Blended Learning

Statistical analysis on the scores of female and male students from the experimental class was conducted to examine students' achievements after the intervention of the "SPOC and Flipped classroom"-based blended teaching strategy and to find out whether there was significant difference between the achievements of male and female students. The results indicated that although male students outperformed female students in their achievements test (m(M) = 78.67, m(F) = 75.04, MD = 75.04), no significant difference was observed ($p > 0.05$). The finding of the data analysis showed that the "SPOC and Flipped classroom"-based blended teaching strategy had no significantly different effects on the achievements of male and female students in their blended learning process. This finding was not consistent with the findings of Al-Haq and Al-Sobh (2010); their study on the effects of a Web-based Writing Instructional EFL Program among Jordanian secondary students' English writing performance showed that there were statistically significant differences ($p < 0.05$) due to gender in favor of female students compared with males [50].

In the early 19th century, linguists found that there were significant differences between men and women in English learning. Since the 1960s, a lot of research has been conducted on the gender differences in sentence patterns, phonemes, and words between men and women and their causes, which has laid a theoretical foundation for English teaching. Although there have been a few studies that have systematically investigated the rate of second language acquisition (SLA) among females versus males, it is a generally accepted fact in language acquisition knowledge that females have a state advantage, initially at least. Some researchers have studied the influence of gender in students' SLAs and have reported that gender-related differences were incidental to their main focus. Farhady (1982) found in his study of 800 university students who were obliged to take a placement test that female students significantly outperformed male students in a listening comprehension test [51]. Eisenstein (1986) also indicated that female students significantly outperformed male studemts in a dialect identification task and in the extent to which they could recognize dialects of greater or lesser prestige [52]. Some other researchers found possible factors for the differences between genders in EFL learning. Ding et al. (2014) pointed that physical factors, learning motivation and will, learning interest, character difference, and learning strategy differences between male and female students were the causes of their different performances in EFL learning [53]. In this case, the finding that male students performed as well as female students in this research could be explained based on the following facts. Firstly, as the intrinsic motivation was slightly stronger, and the English learning motivation was closely related to the English score. The higher the intrinsic motivation, the better the English score [54]. The blended teaching approach implemented in this research has proved to be an effective way of motivating students in EFL learning, since "SPOC and Flipped classroom" were used as two effective teaching strategies in blended teaching to raise students' motivation; therefore, male students' motivation could be stimulated, and their achievements were consequently improved. Thus, the gap of achievements with female students was gradually narrowed down. Secondly, the use of "SPOC and Flipped classroom" techniques gave full play to the function of teacher-guided inquiry teaching mode, in which the integration of online and offline teaching resources, learning activities, teacher–student–peer interactions, and the informative evaluation approach led to a "students-centered" EFL teaching concept. In this way, EFL instruction for the first-year English-major students could follow the principle of "teaching students in accordance with their aptitude", eliminating disadvantages caused by differences in students' learning interests and characters between genders. Therefore, the finding of this research did not find gender: as a moderating variable, to have any significant effect on the differences in achievements between male and female students.

4.5. Regional Background Has Significantly Moderating Effect on the First-Year English-Major Students' Achievements in Their EFL Blended Learning

It has been thought that a student's family background, especially their family's socioeconomic background and their academic performance has a considerable impact. The hypothesis that regional background would be a moderating variable in examining the differences in students' academic performances was derived from the consideration of the different academic performances brought by different information literacy levels among students from urban and rural areas in blended learning. In blended teaching and learning environments, costs should be taken into consideration, since the basic equipment for the online part of blended teaching includes a computer (desktop or laptop), internet access, and other facilities. The researchers supposed that the conditions of a blended teaching environment would be different among students from urban areas and from rural areas; this led to the theory that students from urban areas would have more positive blended learning conditions due to their regional background, that could enable them to better participate in computer-assisted learning or internet learning approaches from a younger age; meanwhile, students from rural areas might not have such good blended learning conditions, and their participation in computer-assisted learning or online learning would have begun at a later age than students from urban areas.

The statistical analysis method of an independent sample *t*-test was carried out on the scores of students from urban areas and from rural areas in the experimental class to verify the hypothesis. The results showed that there was a significant difference between the scores of students from urban areas and those of students from rural areas, and students from urban areas outperformed students from rural areas in their achievements. This finding indicated that the "SPOC and Flipped classroom"-based blended teaching strategy had significantly different effects on the achievements of students from urban areas and students from rural areas, and students from urban areas could obtain better achievements in blended learning. This finding verified the hypothesis that was set in the moderating effects of students' regional backgrounds on their achievements.

Tan and Liu found in their study that a family's economic conditions had a significant impact on a students' academic performance. Both high-income and middle–upper-income students have significantly higher academic performances than low-income students [55]. The reasons were determined to be mainly the following: First of all, the high economic income and the family's living standard could provide a strong material guarantee for children to receive a better education. Secondly, parents from families with poor economic conditions are usually busy making a living and have less time to take care of and supervise their children. In addition, some scholars have shown that a family's economic capital and social capital can help them to choose advantageous educational resources which would, in turn, benefit their children and assist them in achieving better academic results. Hanushek [56] found that the educational level of parents had a significant positive effect on the academic achievement output of their children: the higher the educational level of parents, the better the academic achievement of their children. The International Assessment of Student Ability (PISA) stated that the higher the social and economic status of a student's family was, the better their academic performance was. However, for a long time, scholars have mainly studied students at the basic education stage when analyzing the impact of students' socioeconomic background on their academic performance; few scholars have studied this issue exclusively at the higher education stage. This is especially the case in the new concept of "blended teaching". In conclusion, the differences existing in the students' regional background between urban and rural areas had a significantly moderating effect on the first-year English-major students' achievements of English language skills in the EFL blended teaching using the "SPOC and Flipped classroom"-based blended teaching strategy. This could be attributed to potential reasons of economic levels that could support students' cost of blended learning, parents' educational levels that could influence students' information literacy, or levels of online education development in the regions where students grew up, among other possible reasons.

From the main findings and discussion, in the present study, we have provided referential evidence surrounding the use of the current EFL blended teaching approach for first-year English-major students. The findings here also benefit the understanding of the blended teaching practice of teaching other disciplines. It is important for teachers and educators in different disciplines to reflect on blended teaching during the exploration of achieving optimal results in their approaches to blended teaching through employing appropriate strategies within the scope of different courses, areas, samples, educational levels, and educational policies.

5. Conclusions

The discussion on the findings of the research questions provides pedagogical and theoretical implications for future EFL blended teaching practices among the first-year English-major students. Furthermore, the discussion led to the identification of some important designing principles that needed to be taken into consideration during the development of EFL blended teaching strategies. In this regard, pedagogical implications could be clarified through four perspectives: learners, teacher, course, and technology. Firstly, blended teaching should be designed in accordance with to individual students' differences and learning needs. Secondly, teachers' blended teaching ability and information-based teaching literacy should be improved. Thirdly, it is necessary to enrich online teaching resources and refine curriculum evaluation methods in EFL blended teaching. Finally, it is important to strengthen the techniques used to integrate face-to-face teaching and online instruction.

Because of certain delimitations in the scope of the study, some limitations included the samples participating in the research, the raters who assisted the researcher in conducting the study, and the real-time nature of data collection. The major limitation of the samples was that they were small for both the quasi-experiment and the survey. The samples in the survey were not randomly selected but purposefully selected according to the purpose of study. Additionally, the proportions of the genders and regional backgrounds among students were not balanced. The small size and imbalanced proportions of genders and regional background among the sample shed limit the universal validity and generalizability of the findings. A study with more participants must be conducted to gain more reliable and generalizable outcomes. Secondly, for scoring the tests that students took before and after the quasi-experiment, five raters were selected from among the EFL instructors who were teaching English-major courses at the university. The raters did not receive strict training before rating, but only had a rating standard for each section of the tests. Finally, it is challenging to gain deeper insights into the factors influencing students' achievements from the perspective of a quasi-experiment; the tests did not reflect the causes of the difference in achievements. Therefore, another survey needs to be conducted to explore the influencing factors of students' achievements in blended teaching.

Author Contributions: Conceptualization, L.Z. and K.C.L.; methodology, L.Z.; formal analysis, L.Z.; investigation, L.Z.; data curation, L.Z.; writing—original draft preparation, L.Z.; writing—review and editing, L.Z.; supervision, K.C.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sharma, P.; Barrett, B. *Blended Learning*; Macmillan: Oxford, UK, 2007. Available online: http://biz-e-tech-training.pbworks.com/f/Setting_up_a_wiki_for_an_institution_or_company.pdf (accessed on 1 September 2022).
2. MacDonald, J. *Blended Learning and Online Tutoring*; Gower: Aldershot, UK, 2006; pp. 2–3. Available online: https://www.taylorfrancis.com/books/mono/10.4324/9781315261492/blended-learning-online-tutoring-janet-macdonald (accessed on 20 September 2022).
3. Stracke, E. A road to understanding: A qualitative study into why learners drop out of a blended language learning (BLL) environment. *ReCALL* **2007**, *19*, 57–78. Available online: https://www.cambridge.org/core/journals/recall/article/abs/road-to-understanding-a-qualitative-study-into-why-learners-drop-out-of-a-blended-language-learning-bll-environment/D83CB1A2C49F1E936C56450228A625B8 (accessed on 18 September 2022). [CrossRef]
4. Ulla, M.B.; Perales, W.F. Hybrid teaching: Conceptualization through practice for the post COVID19 pandemic education. *Front. Educ. Front. Media SA* **2022**, *7*, 924594. Available online: https://www.frontiersin.org/articles/10.3389/feduc.2022.924594/full (accessed on 10 October 2022). [CrossRef]
5. Neumeier, P. A closer look at blended learning—Parameters for designing a blended learning environment for language teaching and learning. *ReCALL* **2005**, *17*, 163–178. Available online: https://www.cambridge.org/core/journals/recall/article/abs/closer-look-at-blended-learning-parameters-for-designing-a-blended-learning-environment-for-language-teaching-and-learning/D4CC579CCEA2991DA5078CD0FE53C2AD (accessed on 2 October 2022). [CrossRef]
6. Bilgin, H. Students' CALLing: Blended language learning for students. In *Blended Learning in English Language Teaching: Course Design and Implementation*; British Council: London, UK, 2013; p. 207.
7. Bonk, C.J.; Kim, K.J.; Zeng, T. Future directions of blended learning in higher education and workplace learning settings. In *EdMedia+ Innovate Learning*; Association for the Advancement of Computing in Education (AACE): Chesapeake, VA, USA, 2005; pp. 3644–3649. Available online: https://publicationshare.worldisopen.com/bonk_future.pdf (accessed on 2 October 2022).
8. Wai, C.C.; Seng, E.L.K. Measuring the effectiveness of blended learning environment: A case study in Malaysia. *Educ. Inf. Technol.* **2015**, *20*, 429–443. Available online: https://link.springer.com/article/10.1007/s10639-013-9293-5 (accessed on 15 October 2022). [CrossRef]
9. Siew-Eng, L.; Muuk, M.A. Blended learning in teaching secondary schools' English: A preparation for tertiary Science education in Malaysia. *Procedia-Soc. Behav. Sci.* **2015**, *167*, 293–300. Available online: https://www.sciencedirect.com/science/article/pii/S1877042814068244 (accessed on 22 October 2022). [CrossRef]
10. Harker, M.; Koutsantoni, D. Can it be as effective? Distance versus blended learning in a web-based EAP programme. *ReCALL* **2005**, *17*, 197–216. Available online: https://www.cambridge.org/core/journals/recall/article/abs/can-it-be-as-effective-distance-versus-blended-learning-in-a-webbased-eap-programme/193BC2B9B58EA83C7A8CFDE4316AD36F (accessed on 15 October 2022). [CrossRef]
11. Bai, H.; Zhu, D. A study on mixed teaching mode of Vocational English from the perspective of "Internet+". *J. Zhejiang Polytech. Commun.* **2019**, 54–55. Available online: https://xueshu.baidu.com/usercenter/paper/show?paperid=1e2d0au0hb1x02u0tj6702r031646902&site=xueshu_se (accessed on 10 January 2023).
12. Hu, R. Investigation and Analysis on Learning adaptability of new English majors: A case study of Guizhou Institute of Engineering and Applied Technology. *J. Guizhou Inst. Eng. Appl. Technol.* **2022**, *40*, 120–125. Available online: http://qikan.cqvip.com/Qikan/Article/Detail?id=7106917954 (accessed on 15 November 2022).
13. Xu, M.; Fan, Z.; Gao, J. The helplessness of acquisition in blended teaching for English majors under "Internet + Education". *Educ. Rev.* **2020**, *9*, 106–109. Available online: https://xueshu.baidu.com/usercenter/paper/show?paperid=1h3902b04k7704t0dr4g0eu0xg668758&site=xueshu_se (accessed on 16 November 2022).
14. Sheerah, H.A. Exploring English as Foreign Language (EFL) Students' Perceptions on the Use of Blended Learning to Develop Academic English Language Skills in Preparatory Year in Saudi Arabia. Ph.D. Dissertation, University of Reading, Reading, UK, 2018. Available online: https://centaur.reading.ac.uk/84952/ (accessed on 1 October 2022).
15. Jnr, B.A. An exploratory study on academic staff perception towards blended learning in higher education. *Educ. Inf. Technol.* **2022**, *27*, 3107–3133. Available online: https://link.springer.com/article/10.1007/s10639-021-10705-x (accessed on 10 January 2023).
16. Lin, G.K.; Gong, Q. Research on the Current Situation, Satisfaction and Development of Blended Teaching. *Natl. Acad. For. Grassl. Adm. J.* **2021**, *20*, 43–51+63. Available online: http://s.dic.cool/S/uT4NkmY5 (accessed on 4 February 2023).
17. Yaghmour, K.S. Effectiveness of Blended Teaching Strategy on the Achievement of Third Grade Students in Mathematics. *J. Educ. Pract.* **2016**, *7*, 65–73. Available online: https://eric.ed.gov/?id=EJ1092394 (accessed on 23 December 2022).
18. Saengsawang, P. The Use of Blended Learning to Support Vocabulary Learning and Knowledge Retention in Thai Tertiary EFL Classrooms. Ph.D. Dissertation, Durham University, Durham, UK, 2020. Available online: https://eric.ed.gov/?id=EJ1077202 (accessed on 13 February 2023).

19. Zhao, D.; Jiang, Y. Comparison of English teaching differences between urban and rural areas. *J. Sch. Foreign Lang. Shandong Norm. Univ. (Basic Engl. Educ.)* **2009**, *11*, 47–49. Available online: https://d.wanfangdata.com.cn/periodical/sdsdwgyxyxb200906008 (accessed on 20 February 2023).
20. Peng, B. Analysis on the Causes and Countermeasures of Urban and Rural Differences Affecting Students' English Learning. Extralinguistic Learning: High School Edition. 2014. p. 1. Available online: https://d.wanfangdata.com.cn/periodical/yswxx-yyjy201403021 (accessed on 14 December 2022).
21. Lv, Y. College Students' English Learning Concepts and Their Influencing Factors. Ph.D. Dissertation, Henan University, Kaifeng, China, 2008.
22. Han, W.; Zhang, X.; Lu, T.; Sun, Y.I. Spoc-based mixed teaching model curriculum construction: A case study of "Software Project Management" course. *Ind. Inf. Educ.* **2015**, *48*, 38–42. Available online: http://qikan.cqvip.com/Qikan/Article/Detail?id=666731811 (accessed on 22 September 2022).
23. Lage, M.J.; Platt, G.J.; Treglia, M. Inverting the classroom: A gateway to creating an inclusive learning environment. *J. Econ. Educ.* **2015**, *31*, 30–43. Available online: http://www.maktabe-hekmat.ir/wp-content/uploads/2018/12/2000-lage.pdf (accessed on 6 January 2023). [CrossRef]
24. Baker, W.J. The "classroom flip": Using web course management tools to become the guide by the side. In Proceedings of the 1st Annual Higher Education Flipped Learning Conference, Greeley, CO, USA, 8–10 June 2016; pp. 15–24. Available online: https://digscholarship.unco.edu/cgi/viewcontent.cgi?article=1058&context=heflc#page=21 (accessed on 8 June 2023).
25. Bishop, J.; Verleger, M.A. The flipped classroom: A survey of the research. In Proceedings of the 2013 ASEE Annual Conference & Exposition, Atlanta, Georgia, 23–26 June 2013; pp. 23–1200. Available online: https://peer.asee.org/the-flipped-classroom-a-survey-of-the-research (accessed on 17 March 2023).
26. Abeysekera, L.; Dawson, P. Motivation and cognitive load in the flipped classroom: Definition, rationale and a call for research. *High. Educ. Res. Dev.* **2015**, *34*, 1–14. Available online: https://www.tandfonline.com/doi/abs/10.1080/07294360.2014.934336 (accessed on 23 January 2023). [CrossRef]
27. Peers, I. *Statistical Analysis for Education and Psychology Researchers: Tools for Researchers in Education and Psychology*; Routledge: Abingdon, UK, 2006. Available online: https://cds.cern.ch/record/996878 (accessed on 15 February 2023).
28. Du, X. *Experimental Design and SPSS Data Processing in Psychology and Education Research*; Peking University Press: Beijing, China, 2013. Available online: http://hn.sslibrary.com/showbook.do?dxNumber=13236593&d=4EDFCECBF2E46AB36F5AA02009350935&fFenleiID=0G404005 (accessed on 28 November 2022).
29. West, S.G.; Finch, J.F.; Curran, P.J. Structural Equation Models with Nonnormal Variables: Problems and Remedies. In *Structural Equation Modeling: Issues, Concepts, and Applications*; Hoyle, R.H., Ed.; Sage: Newbury Park, CA, USA, 1995; pp. 56–75. Available online: https://psycnet.apa.org/record/1995-97753-004 (accessed on 13 March 2023).
30. Kim, H.Y. Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis. *Restor. Dent. Endod.* **2013**, *38*, 52–54. Available online: https://synapse.koreamed.org/articles/1090037 (accessed on 4 April 2023). [CrossRef]
31. Oweis, T.I. Effects of using a blended learning method on students' achievement and motivation to learn English in Jordan: A pilot case study. *Educ. Res. Int.* **2018**, *2018*, 7425924. Available online: https://www.hindawi.com/journals/edri/2018/7425924/ (accessed on 8 April 2022). [CrossRef]
32. Bañados, E. A blended-learning pedagogical model for teaching and learning EFL successfully through an online interactive multimedia environment. *CALICO J.* **2006**, 533–550. Available online: https://www.jstor.org/stable/24156354 (accessed on 4 February 2022). [CrossRef]
33. Aji, M.P.P. English listening blended learning: The implementation of blended learning in teaching listening to university students. *Kaji. Linguist. Dan Sastra* **2017**, *2*, 25–32. Available online: https://journals.ums.ac.id/index.php/KLS/article/view/5349 (accessed on 14 June 2023). [CrossRef]
34. Bolandifar, S. Effects of blended learning on reading comprehension and critical thinking skills of undergraduate ESL students. Malaysia: University Putra Malaysia. Ph.D. Dissertation, University Putra Malaysia, Serdang, Malaysia, 2017. Available online: http://psasir.upm.edu.my/id/eprint/69634/1/fpp%202017%206%20ir.pdf (accessed on 20 March 2023).
35. Ghazizadeh, T.; Fatemipour, H. The effect of blended learning on EFL learners' reading proficiency. *J. Lang. Teach. Res.* **2017**, *8*, 606. Available online: http://academypublication.com/issues2/jltr/vol08/03/21.pdf (accessed on 16 November 2022). [CrossRef]
36. Yudhana, S. The Implementation of Blended Learning to Enhance English Reading Skills of Thai Undergraduate Students. *Engl. Lang. Teach.* **2021**, *14*, 1–7. Available online: https://eric.ed.gov/?id=EJ1308909 (accessed on 22 December 2022). [CrossRef]
37. Hamouda, A. The effect of blended learning on developing Saudi English majors' writing skills. *Int. J. Engl. Educ.* **2018**, *7*, 40–83. Available online: https://ijee.org/assets/docs/4ARAFAT.9883315.pdf (accessed on 24 April 2023).
38. Ginaya, G.; Rejeki IN, M.; Astuti NN, S. The effects of blended learning to students' speaking ability: A study of utilizing technology to strengthen the conventional instruction. *Int. J. Linguist. Lit. Cult.* **2018**, *4*, 1–14. Available online: https://www.researchgate.net/profile/Gede-Ginaya/publication/324905956 (accessed on 28 March 2023).
39. Benhadj, Y. A Quasi-experimental Study on the Impact of BlendedLearning on EFL Students' Language Proficiency. *Int. J. Lang. Lit. Stud.* **2021**, *3*, 146–155. Available online: https://www.ijlls.org/index.php/ijlls/article/view/668 (accessed on 4 June 2023). [CrossRef]

40. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Lawrence Erlbaum Associates: Mahwah, NJ, USA, 1988; pp. 8–13. Available online: https://www.tandfonline.com/doi/abs/10.1080/00401706.1989.10488618 (accessed on 24 December 2022).
41. Boelens, R.; De Wever, B.; Voet, M. Four key challenges to the design of blended learning: A systematic literature review. *Educ. Res. Rev.* **2017**, *22*, 1–18. Available online: https://www.sciencedirect.com/science/article/abs/pii/S1747938X17300258 (accessed on 25 February 2023). [CrossRef]
42. Alnoori, B.; Obaid, S. The Effectiveness of 50-50 Blended Learning Method on Teaching Reading Skills in ESL Classroom. *Int. J. Lang. Acad.* **2017**, *5*, 288–303. Available online: https://www.researchgate.net/profile/Bushra-Alnoori/publication/321971507 (accessed on 4 March 2023). [CrossRef]
43. Kirkgoz, Y. A Blended Learning Study on Implementing Video Recorded Speaking Tasks in Task-Based Classroom Instruction. *Turk. Online J. Educ. Technol.-TOJET* **2013**, *10*, 1–13. Available online: https://eric.ed.gov/?id=EJ946605 (accessed on 24 February 2023).
44. Yang, F.; Wei, X.; Zhang, W. An Analysis of Blended Teaching Mode of College English. *Foreign Lang. Teach.* **2017**, *12*, 21–28. Available online: https://d.wanfangdata.com.cn/periodical/wydhjx201701004 (accessed on 15 September 2022).
45. Hung, H.T. Flipping the classroom for English language learners to foster active learning. *Comput. Assist. Lang. Learn.* **2015**, *28*, 81–96. Available online: https://d.wanfangdata.com.cn/periodical/5b6838bc948fbd45354cbe2a04aad09a (accessed on 21 January 2023). [CrossRef]
46. Alshumaimeri, Y.A.; Almasri, M.M. The effects of using WebQuests on reading comprehension performance of Saudi EFL students. *TOJET Turk. Online J. Educ. Technol.* **2012**, *11*, 295–306.
47. Tang, W.L. An investigation and analysis of motivational factors affecting oral English learning. *Foreign Lang. Teach. Learn.* **2005**, *26*, 65–68. Available online: https://xueshu.baidu.com/usercenter/paper/show?paperid=120r0mk0xb2c0g20hg6d06u0tk150764&site=xueshu_se (accessed on 19 February 2023).
48. Wang, Y. On the improvement of English majors' oral competence. *J. Hotan Norm. Coll. Chin. Compr. Ed.* **2007**, *27*, 151–152.
49. Zhang, N. The practical application of blended learning in oral English teaching. *J. Kaifeng Inst. Educ.* **2017**, *37*, 90–91. Available online: https://d.wanfangdata.com.cn/periodical/kfjyxyxb201707041 (accessed on 22 April 2023).
50. Al-Haq, F.; Al-Sobh, M.A. The effect of a web-based writing instructional EFL program on enhancing the performance of Jordanian secondary students. *JALT Call J.* **2010**, *6*, 189–218. Available online: https://www.researchgate.net/publication/340942498 (accessed on 14 March 2023). [CrossRef]
51. Farhady, H. Measures of Language Proficiency from the Learner's Perspective. *TESOL Q.* **1982**, *16*, 43–59. Available online: https://onlinelibrary.wiley.com/doi/abs/10.2307/3586562 (accessed on 24 February 2023). [CrossRef]
52. Eisenstein, M. A study of Social Variation in Adult Second Language Acquisition. *Lang. Learn.* **1982**, *32*, 367–391. Available online: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-1770.1982.tb00977.x (accessed on 24 February 2023). [CrossRef]
53. Ding, J.; Lai, Y.; Yu, W. A study on the difference in English Learning ability between male and female non-English majors in independent colleges. *Engl. Sq. (Acad. Res.)* **2014**, 120–121. Available online: http://www.cqvip.com/QK/61348X/20141/48286735.html (accessed on 13 November 2022).
54. Lu, Y. A study on the correlation between English learning motivation and academic achievement. *Educ. Teach. Forum* **2015**, 2. Available online: http://www.cqvip.com/QK/60535A/201542/666196805.html (accessed on 5 January 2023).
55. Tan, Y.; Liu, Z.C. A study on the relationship between family socioeconomic background and academic achievement of college students. *Exam. Wkly.* **2012**, 3. Available online: https://d.wanfangdata.com.cn/periodical/kszk201214132 (accessed on 24 April 2023).
56. Hanushek, E.A. What matters for student achievement. *Educ. Next* **2016**, *16*, 18–26. Available online: http://www.k12accountability.org/resources/At-Risk-Students/hanushek_on_coleman.pdf (accessed on 12 February 2023).

sustainability

Review

Transforming Education: A Comprehensive Review of Generative Artificial Intelligence in Educational Settings through Bibliometric and Content Analysis

Zied Bahroun [1,*], Chiraz Anane [2], Vian Ahmed [1] and Andrew Zacca [1]

1 Department of Industrial Engineering, College of Engineering, American University of Sharjah, Sharjah P.O. Box 26666, United Arab Emirates; vahmed@aus.edu (V.A.); b00084411@aus.edu (A.Z.)
2 Department of Foreign Languages, College of Arts, Humanities and Social Sciences, University of Sharjah, Sharjah P.O. Box 27272, United Arab Emirates; canane@sharjah.ac.ae
* Correspondence: zbahroun@aus.edu

Abstract: In the ever-evolving era of technological advancements, generative artificial intelligence (GAI) emerges as a transformative force, revolutionizing education. This review paper, guided by the PRISMA framework, presents a comprehensive analysis of GAI in education, synthesizing key insights from a selection of 207 research papers to identify research gaps and future directions in the field. This study begins with a content analysis that explores GAI's transformative impact in specific educational domains, including medical education and engineering education. The versatile applications of GAI encompass assessment, personalized learning support, and intelligent tutoring systems. Ethical considerations, interdisciplinary collaboration, and responsible technology use are highlighted, emphasizing the need for transparent GAI models and addressing biases. Subsequently, a bibliometric analysis of GAI in education is conducted, examining prominent AI tools, research focus, geographic distribution, and interdisciplinary collaboration. ChatGPT emerges as a dominant GAI tool, and the analysis reveals significant and exponential growth in GAI research in 2023. Moreover, this paper identifies promising future research directions, such as GAI-enhanced curriculum design and longitudinal studies tracking its long-term impact on learning outcomes. These findings provide a comprehensive understanding of GAI's potential in reshaping education and offer valuable insights to researchers, educators, and policymakers interested in the intersection of GAI and education.

Keywords: education; generative artificial intelligence; educational technology; ChatGPT; ethics; review

Citation: Bahroun, Z.; Anane, C.; Ahmed, V.; Zacca, A. Transforming Education: A Comprehensive Review of Generative Artificial Intelligence in Educational Settings through Bibliometric and Content Analysis. *Sustainability* **2023**, *15*, 12983. https://doi.org/10.3390/su151712983

Academic Editors: Savvas A. Chatzichristofis, Hao-Chiang Koong Lin, Jian-Hong Ye, Yung-Wei Hao and Yu-Feng Wu

Received: 1 August 2023
Revised: 17 August 2023
Accepted: 17 August 2023
Published: 29 August 2023

1. Introduction

In an era where technological advancements are redefining human existence, there is one revolutionary force that stands out among the rest: artificial intelligence (AI). AI has permeated every corner of our lives, quietly and profoundly transforming the way we work, communicate, and navigate the world. The rise of AI was sparked in November 2022, when U.S. company OpenAI (San Francisco, CA, USA) released ChatGPT [1], an AI program that draws upon a large language database to generate responses from text-based inputs entered by humans [2]. Since its launch, it has skyrocketed in popularity, becoming one of the fastest-growing consumer applications ever with an estimated 100 million active users every month [3]. Following the successful release of ChatGPT, other tech companies have come out with their own AI-powered products, such as Google's Bard [4] and GitHub's Copilot [5].

AI has infused various domains of our lives, revolutionizing industries such as finance, education, engineering, healthcare, and many more. According to Hadi et al. [6], AI has transformed practices in the financial sector, such as algorithmic trading, market prediction, and financial reporting. Hadi et al. [6] further argue that education has seen the integration of AI in automated essay grading, intelligent tutoring systems, and personalized learning. In engineering, AI drives advancements in code generation, software testing, and document generation. In addition, in the healthcare sector, AI assists in radiologic decision-making, patient care, and medical education. These various use cases demonstrate the extensive reach and transformative power that AI can deliver across these diverse domains.

More notably, the field of education is experiencing a transformative revolution, driven by the emergence of GAI. According to Baidoo-Anu and Owusu Ansah [7], GAI is an unsupervised or partially supervised machine learning framework that seamlessly generates artificial creations by analyzing existing digital content like videos, images/graphics, text, and audio. The impact of GAI has been felt across a diverse range of educational fields. In education, GAI finds versatile applications in assessment and evaluation, student performance prediction, intelligent tutoring systems, and learning management [8]. In the medical field, the application of GAI, such as ChatGPT, extends to medical education, where it offers realistic case scenarios and instantaneous feedback to medical students on their diagnostic and treatment decisions [9]. Moreover, GAI proves instrumental in providing simplified explanations of complex medical and pharmaceutical concepts, facilitating students' understanding [9,10]. In the field of computer science, AI tools demonstrate tremendous potential in coding education, as evidenced by GPT-3's ability to generate diverse code explanations [11].

AI's impact has triggered a wide array of discussions encompassing diverse topics, ranging from its potential in transforming learning and teaching methods [12] to its role in aiding research endeavors and to the crucial considerations of ethics and academic integrity in its implementation [13]. In the domain of teaching, GAI showcases promising opportunities for lesson planning, personalized learning support, rapid assessment and evaluation, and addressing learners' queries [12]. Furthermore, Markel et al. [14] have introduced a novel AI tool called "GPTeach", which is specifically designed for teacher training, enabling aspiring educators to practice teaching with simulated students powered by GPT. As the transformative influence of AI in education continues to unfold, these discussions will undoubtedly shape the future of how we impart knowledge and foster learning in an ever-evolving educational landscape.

This transformative technology has opened up new avenues for personalized instruction, enhanced feedback, and tailored learning experiences, ultimately paving the way for a more efficient, inclusive, and engaging educational environment. GAI is reshaping the way we teach and learn by offering innovative solutions to longstanding challenges and opening up new possibilities for improving educational experiences for learners of all ages. As GAI continues to advance, it is crucial to investigate its applications, implications, and the challenges it poses to further explore its role in shaping the future of education.

The scope of this review encompasses a wide range of topics, delving into the applications of GAI in various educational contexts, the tools and techniques utilized, the effectiveness of GAI in supporting teaching and learning, the impact on student outcomes, and the potential challenges and ethical considerations associated with its implementation. By examining multiple dimensions of GAI in education, this review aims to provide a holistic overview of the field. This paper aims to analyze and consolidate a compilation of existing reviews on GAI in specific fields of education. By thoroughly examining a diverse range of papers in this domain, this study offers a comprehensive and in-depth understanding of the current state of research in this field. Through synthesizing and summarizing key findings, methodologies, and recommendations from multiple papers, this study provides a valuable resource for researchers, educators, and policymakers who are keenly interested in exploring the transformative intersection of GAI and education.

This review paper contributes to the field in various ways. Firstly, it provides a comprehensive overview of the current state of research on GAI in education, allowing researchers to identify the prevailing themes and research directions within the field. Secondly, it synthesizes the findings and insights from multiple review papers, providing a holistic perspective on the effectiveness and potential of GAI in educational contexts. Additionally, this review identifies research gaps and areas that require further investigation, guiding future research endeavors. The subsequent sections of this paper will, therefore, present the methodology employed in selecting and analyzing the research papers, followed by a comprehensive content analysis and synthesis of the findings and insights. Key themes, trends, and recommendations will be presented, allowing readers to gain a comprehensive understanding of the existing knowledge base. Subsequently, a comprehensive bibliometric analysis will be conducted, examining prominent AI tools, research focus, geographic distribution, and interdisciplinary collaboration in detail. Finally, this review will conclude by highlighting the implications of the synthesized findings and suggesting future research directions to advance the field of GAI in education.

2. Methodology

The methodology employed in this research follows a systematic approach to gather and analyze literature concerning the integration of GAI in education. It involves four key steps, as described below:

a. Literature retrieval—This step entails selecting appropriate search terms and keywords to comprehensively capture pertinent publications related to the intended topic representing the initial and pivotal phase of the data collection process. We retrieved a compilation of pre-existing articles and publications in the domain of GAI in education from the *Scopus* database. By utilizing a combination of the following keywords, "Education", "GAI", "Generative Artificial Intelligence", "GPT-4", "ChatGPT", "AlphaCode", "GitHub Copilot", and "Bard", the authors conducted a targeted search across titles, abstracts, and keyword fields. This process resulted in the enlistment of a total of 437 papers, spanning the years 2018 to 2023.
b. Literature screening—The literature screening process for this study drew inspiration from the PRISMA statement, which is widely recognized as a rigorous and transparent approach for conducting systematic reviews and meta-analyses (Figure 1). PRISMA provides a structured framework to ensure the systematic identification, selection, and evaluation of the relevant literature, promoting the reliability and reproducibility of the review process [15]. Initially, a total of 437 papers were retrieved. Following the removal of duplicates, 312 papers remained. After observingly reviewing each publication, all articles that were deemed irrelevant to our research focus were discarded. We narrowed the list down to 217 reviews, papers, and publications that spanned from the years 2018 to 2023. The chronological growth of the papers (in the context of the topic) in the above-mentioned spanned years is depicted in Figure 2.

The figure above depicts a substantial number of papers available from our compiled list, specifically published in the year 2023. Notably, there is a significant and exponential growth in the field of GAI from 2018 to 2023, indicating a growing amount of research interest in the topic. This surge in research can be attributed to the recent emergence of popular GAI tools, such as ChatGPT, which have garnered considerable attention and sparked innovation within the education domain. The increasing research activity, therefore, suggests that GAI is perceived as a promising area for innovation in education and that researchers and educators are increasingly recognizing the potential benefits of integrating AI technologies into educational settings.

c. Content analysis—This step entails carefully studying a large amount of information, like research publications, and organizing it in a meaningful way. This helps researchers see common themes and patterns in the information. In this case, the content analysis focused on how GAI is being used in education. The methodology involved categorizing and classifying the research papers into different themes and sub-themes. This organization helps researchers understand the different ways GAI is impacting education, making it easier to draw conclusions and insights from the collected information.

d. Bibliometric analysis—It serves the purpose of systematically evaluating academic literature, typically through analyzing citations and references in research papers. This methodology helps researchers understand the influence, trends, and relationships among different academic works. By examining patterns of citations, co-authorships, and keywords, bibliometric analysis allows for the identification of key authors, influential papers, emerging areas of study, and collaborative networks within a particular field. This approach aids researchers in gaining insights into the evolution of research topics, identifying leading scholars.

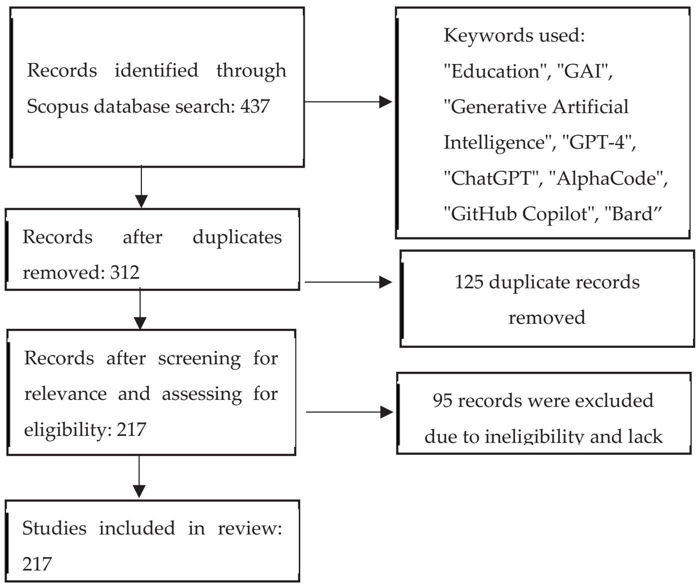


Figure 1. Literature screening approach.

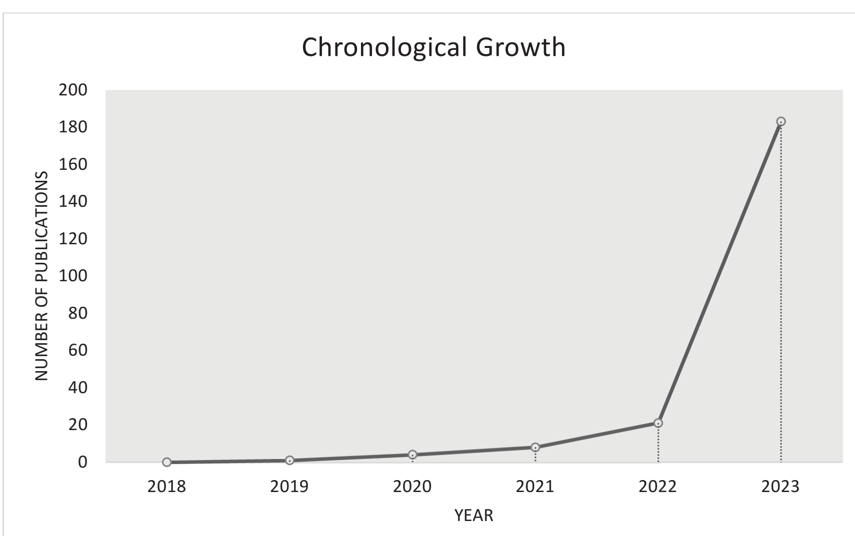


Figure 2. The number of papers and publications in relation to GAI in education (2018–2023).

The remaining part of this paper presents a content analysis and a bibliometric analysis while adhering to the methodological steps described in this section.

3. Content Analysis

The integration of GAI in education has become a subject of considerable interest, opening up new possibilities and challenges for educational practices. GAI models, such as ChatGPT, have been extensively explored across various academic disciplines, presenting transformative opportunities in teaching, learning, and research. This section seeks to provide a thorough content analysis of the use of GAI in education.

3.1. Application of GAI in Computer Science

This study explores a collection of research papers focusing on the intersection of GAI and computer science education and assesses various aspects of GAI technologies and their implications for teaching and learning programming concepts (Table 1).

Table 1. The applications of GAI in computer science.

Themes	Authors	Focus
Implications and potential of GAI code generation tools in programming education	Jonsson and Tholander (2022) [16]	Highlights the transformative potential of generative machine learning in creative programming education, fostering both programming proficiency and creativity through students' exploration of the tool's inconsistent behavior.
	Finnie-Ansley et al. (2022) [17]	Investigates the performance and implications of Codex, a deep learning model, in introductory programming, showcasing its ability to outperform students in problem-solving while raising considerations for computing education.
	MacNeil et al. (2022) [11]	Explores the potential of large language models (LLMs) in providing explanations for programming concepts, contributing to the understanding of LLMs' role in enhancing learning experiences.
	Denny et al. (2023) [18]	Investigates Copilot, an AI-powered code generation tool, emphasizing its problem-solving capabilities and discussing prompt engineering as a means to enhance computational thinking skills.
	Kazemitabaar et al. (2023) [19]	Demonstrates the positive influence of AI code generators, like OpenAI Codex, on novice programmers, highlighting improvements in performance, computational thinking skills, and motivation.
Exploration and utilization of AI tools for programming education	Savelka et al. (2023) [20]	Investigates the effectiveness of generative pre-trained transformer (GPT) models in answering code-related multiple-choice questions, highlighting the challenges faced by GPT models in code analysis and reasoning.
	Wermelinger (2023) [21]	Focuses on GitHub Copilot, an AI-powered programming assistant, evaluating its performance and limitations in code generation, explanation provision, test generation, and bug fixing.

Table 1. *Cont.*

Themes	Authors	Focus
Ethical considerations and pedagogical approaches in AI education	Becker et al. (2023) [22]	Emphasizes the need for collaborative efforts to explore the implications and effectiveness of AI-driven code generation tools in computing education, highlighting the importance of shaping their trajectory.
	Jacques (2023) [23]	Challenges the notion that AI-generated coding tools render human programming obsolete, advocating for a reevaluation of teaching methods that foster problem-solving, critical thinking, and communication skills.
	Raji et al. (2021) [24]	Delves into the ethical considerations in AI education and proposes a shift toward a collaborative and holistic pedagogy that bridges the gap between computer science and humanities and social sciences.
Integration of AI technologies in programming education	Yilmaz and Yilmaz (2023) [25]	Explores the incorporation of ChatGPT into programming education and highlights its positive impact on computational thinking skills, programming self-efficacy, and student motivation.
	Cortinas-Lorenzo and Lacey (2023) [26]	Addresses the need for transparency and interpretability in affective computing systems, highlighting the challenges associated with multimodal data, context integration, and interaction capturing.

The findings from the above publications reveal several key insights. Firstly, researchers highlight the transformative potential of generative machine learning in creative programming education, indicating that GAI can significantly impact how programming concepts are taught and learned. Secondly, crucial considerations are raised regarding AI-generated code, pointing to the need for careful examination of the quality and reliability of code produced by AI systems. Moreover, the potential of large language models (LLMs) in providing explanations for programming concepts is showcased, suggesting that AI can enhance the understanding of complex programming principles. Additionally, practical insights are offered into the implications of AI-powered code generation tools, like Copilot, shedding light on the benefits and challenges of using such technologies in educational settings.

These findings, therefore, provide valuable insights into GAI in computer science education, highlighting a lack of in-depth exploration of drawbacks and ethical concerns. While the spotlight is often on the positive outcomes of AI-generated code, it is crucial to acknowledge the possible drawbacks, like an excessive dependence on automated coding and its potential to undermine students' algorithmic comprehension. Furthermore, the studies focus on short-term improvements in performance and motivation without fully considering long-term effects on problem-solving skills and holistic learning. Future research should, therefore, balance the positive and negative aspects of GAI integration, addressing academic integrity, assessment challenges, and innovative methods for incorporating AI-generated code into programming curricula.

3.2. Application of GAI in Engineering Education

This section delves into a collection of research papers that explore the impact of GAI in various disciplines within engineering education. The papers are, therefore, categorized into three groups, as shown in Table 2.

Table 2. The application of GAI in engineering education.

Themes	Authors	Focus
Transformative role of AI in education and learning	Marquez et al. (2023) [27]	Emphasizes the potential of integrating product-based learning strategies and AI text generation models in biobased materials education, showcasing how this approach promotes active participation, problem-solving skills, and the generation of sustainable solutions.
	Humphry and Fuller (2023) [28]	Demonstrates the utility of ChatGPT in revolutionizing chemistry education by effectively generating discussion sections for laboratory reports, highlighting its potential to improve student learning experiences and streamline the report writing process.
	Sánchez-Ruiz et al. (2023) [29]	Investigates ChatGPT's impact on blended learning in engineering education, particularly mathematics. The study reveals that the integration of ChatGPT positively influences students' critical thinking abilities, problem-solving skills, and collaborative work. However, it also raises concerns about potential effects on the development of lateral competencies, highlighting the need for adaptive teaching strategies in blended learning environments.
	Nikolic et al. (2023) [30]	Analyzes the impact of ChatGPT on assessment practices in engineering education.
AI in design and collaboration	Gmeiner et al. (2023) [31]	Identifies challenges faced by designers when collaborating with AI-based design tools, including difficulties in understanding and adjusting AI-generated outputs and effectively communicating design objectives.
	Chen et al. (2020) [32]	Sheds light on the transformative influence of AI in engineering design education, enabling the realization of previously unattainable solutions and concepts.
	Galdon et al. (2021) [33]	Presents an operational framework that harnesses abductive reasoning and probabilistic knowledge to enrich design and engineering education. The study showcases the framework's potential in fostering creative engagement with knowledge and generating impactful solutions, emphasizing the necessity of reframing probabilistic knowledge and embracing transformative approaches to address real-world uncertainties and challenges.
Innovative applications of social robots	Elfaki et al. (2023) [34]	Introduces a cloud-based framework aimed at augmenting the intelligence and autonomy of social robots in engineering education. The study showcases the feasibility and potential of this framework, empowering social robots to showcase heightened capabilities through the utilization of cloud computing and clustering.

GAI presents transformative benefits in engineering education, as evident from various studies. Indeed, its integration enhances active learning and problem-solving skills and streamlines educational processes through advanced chatbots and text generation models. The potential to foster creative engagement in design and collaboration further empowers students to generate innovative solutions. Cloud-based frameworks amplify social robots' intelligence and accessibility in engineering education, providing new opportunities for interactive learning experiences.

Taking a critical view of these studies, while they undoubtedly showcase the positive implications of GAI for fostering active learning, honing problem-solving skills, and streamlining design processes, they also unveil crucial issues that demand careful consideration. These concerns encompass the potential erosion of lateral competencies due to overreliance on GAI, challenges in human–AI collaboration, and the necessity for a pedagogical framework that seamlessly incorporates GAI tools without diluting the essence of education. While these research papers accentuate the transformative potential of GAI in revolutionizing the educational landscape, they simultaneously underscore the urgency for judicious implementation, particularly within dynamic domains, like design and social robots. To address these issues, future work must refine GAI integration and assess long-term impacts to ensure its benefits are maximized while mitigating potential challenges.

3.3. Application of GAI in Higher Education

This section explores the potential, applications, challenges, and implications of GAI tools in specifically higher education. Table 3 shows the classification of these papers and their main focus.

Table 3. The application of GAI in higher education.

Publications	Authors	Focus
The integration of GAI tools such as ChatGPT in higher education	Konecki et al. (2023) [35]	Conducts a comprehensive examination of ChatGPT's role as an educational aid, providing insights into its benefits and limitations.
	Eager and Brunton (2023) [36]	Focuses on the broader implications of LLMs and conversational-style GAI on pedagogy. They emphasize the importance of implementing effective plagiarism detection mechanisms, fostering supportive learning environments, and offering faculty development programs to ensure the responsible integration of AI in teaching and learning.
	Chaudhry et al. (2023) [37]	Intensively examines the impact of ChatGPT on learning outcomes and academic integrity in undergraduate degree programs. The study evaluates ChatGPT's performance in handling assignments and compares it to high-achieving students. Additionally, the effectiveness of plagiarism detection tools is assessed in the context of ChatGPT usage. The findings highlight the limitations of ChatGPT in its ability to fully replace human performance and underscore the need for a reevaluation of student evaluation approaches in light of AI-based tools.
Students' acceptance and use of GAI technology in higher education	Iskender (2023) [38]	Addresses the ethical and pedagogical considerations of GAI in higher education and conducted an interview with ChatGPT to delve into its impact on the realm of higher education and academic publishing. The interview sheds light on ChatGPT's potential in grading and supporting students, acknowledging its usefulness in certain educational contexts. However, it also raises crucial concerns about potential risks, such as the possibility of diminishing critical thinking skills and exacerbating educational inequalities.
	Strzelecki (2023) [39]	Delves into the adoption and use of ChatGPT among higher education students, developing a model based on technology adoption theory. The study identifies key predictors such as habit, performance expectancy, and hedonic motivation in determining students' acceptance and usage of ChatGPT.

Table 3. *Cont.*

Publications	Authors	Focus
	Stojanov (2023) [40]	Presents an autoethnographic study exploring the utilization of ChatGPT as a learning aid. The study underscores the potential benefits of ChatGPT while acknowledging its limitations, including occasional superficiality and inconsistency in generated answers. It emphasizes the need for caution and further research to fully understand the implications of incorporating ChatGPT in educational contexts.
The applications of AI and ML in higher education	Pinto et al. (2023) [41]	Conducts a systematic literature review to shed light on the diverse uses of ML in higher eduction. The review highlights the prediction of academic performance and employability as the most extensively studied application of ML in higher education.

The studies in this section offer insights into various aspects of GAI integration in higher education. They delve into the integration of GAI tools, such as ChatGPT and large language models, exploring their potential impact on teaching and learning practices. Student acceptance and usage of GAI technology are examined, revealing factors that influence their adoption. The applications of AI and machine learning in higher education are explored, particularly in predicting academic performance and employability. Furthermore, the impact of ChatGPT on learning outcomes and academic integrity is assessed, prompting discussions on assessment methodologies. Ethical and pedagogical considerations of GAI are discussed, highlighting both potential benefits and concerns related to originality and critical thinking.

These studies provide valuable insights into the potential of GAI tools in enhancing higher education. However, there are certain areas that could be more comprehensive, whereby the tendency to highlight the positive aspects of GAI integration could benefit from a more balanced exploration of potential challenges and unintended consequences. For instance, while the impact of tools like ChatGPT on student learning is explored, a deeper investigation into how such tools might affect critical thinking skills and originality in academic work could provide a more holistic understanding. Additionally, while the studies touch on students' acceptance of GAI technology, a deeper analysis of the socio-cultural factors influencing students' attitudes and perceptions is warranted. By addressing these critical perspectives, future research in this area can provide a more nuanced and comprehensive understanding of the role and impact of GAI in higher education.

3.4. Application of GAI in Medical Education

This section aims to provide an overview of research papers that discuss the integration of GAI tools in healthcare, as shown in Table 4.

Table 4. The application of GAI in medical education.

Publications	Authors	Focus
The integration of GAI tools in healthcare	Sallam et al. (2023) [9]	Investigates the potential benefits and limitations of incorporating ChatGPT into medical fields. The study reveals the promising advantages of ChatGPT, such as personalized learning, clinical reasoning support, and better comprehension of complex concepts in medical education. It highlights concerns related to data privacy, potential bias in generated content, and possible negative effects on critical thinking and communication skills.

Table 4. *Cont.*

Publications	Authors	Focus
	Waisberg et al. (2023) [42]	Investigates the potential applications of GPT-4 in neurosurgery, revealing its ability to enhance precision and effectiveness in neurosurgical practices, leading to improved patient outcomes and advancements in the field.
	Sevgi et al. (2023) [43]	Assesses the reliability of ChatGPT in neurosurgical education, pointing out that while it provides interesting responses, its credibility is compromised by the lack of citations for scientific queries, suggesting caution in relying solely on ChatGPT as an educational resource.
	Šlapeta (2023) [44]	Explores the advantages of integrating large language models, including ChatGPT, in parasitology education, highlighting their potential to improve learning outcomes and facilitate knowledge acquisition in the field.
	Huh (2023) [45]	Compares ChatGPT's knowledge and interpretive abilities with those of medical students in Korea in parasitology, revealing its inferior performance and indicating the need for further refinement in specialized medical contexts.
	Dolezal et al. (2023) [46]	Investigates the use of deep neural networks and conditional generative adversarial networks (cGANs) to generate synthetic histology for tumor classification, showcasing its value as an educational tool for enhancing understanding of tumor biology.
	Skalidis et al. (2023) [47]	Examines ChatGPT's performance in answering questions from a post-graduate exam in core cardiology, demonstrating its impressive success in handling higher-level examination content and raising discussions on its potential applications in diverse educational contexts within cardiology.
The applications of GAI in clinical education and practices	Corsello and Santangelo (2023) [48]	Discusses the potential impact of ChatGPT on future pediatric research, highlighting positive effects such as improved clinical decision-making and enhanced medical education. However, the study also addresses concerns about bias, fairness, and ethical considerations associated with GAI utilization.
	Lyu et al. (2023) [49]	Investigates the feasibility of utilizing ChatGPT for translating radiology reports into plain language, revealing its effective translation capabilities while noting occasional occurrences of random or oversimplified responses that require attention.
	Abdel-Messih and Kamel Boulos (2023) [50]	Demonstrates ChatGPT's ability to successfully pass the United States Medical Licensing Examination (USMLE) by effectively answering queries related to clinical toxicology, showcasing its potential in this domain.
	Oh et al. (2023) [51]	Explores the feasibility of integrating ChatGPT into surgical education and training, acknowledging its successful translation outcomes in radiology reports while also recognizing the need to address occasional occurrences of random or oversimplified responses.
	Lee et al. (2023) [52]	Discusses the implications of GPT-4 for a chatbot specifically tailored for the field of medicine. The study examines the potential impacts of this well-educated chatbot on the practice of medicine.

Table 4. *Cont.*

Publications	Authors	Focus
The potential of GAI in health education and social interaction	Yang et al. (2023) [53]	Focuses on the utilization of ChatGPT in assisting children with Down syndrome, conducting a literature review to explore its advantages and limitations in various domains for individuals with this condition. It highlights the significance of ChatGPT as a valuable tool for promoting cognitive development, education, and social interaction, catering to the unique requirements of children with Down syndrome.
	Wang et al. (2023) [54]	Explores the role of GAI in pandemic response and future medicine and conducts a comprehensive review of the potential applications of AI and GAI in the prevention and treatment of the COVID-19 pandemic.
The ethical considerations and responsible use of GAI in healthcare	Cascella et al. (2023) [55]	Investigates the potential applications and limitations of ChatGPT in healthcare, emphasizing the need to educate stakeholders about the appropriate utilization of AI-based language models in medicine while addressing the ethical challenges they may present.
	Gupta et al. (2023) [56]	Examines the use of ChatGPT in generating innovative ideas in plastic surgery, highlighting its potential to contribute to research and support various aspects of patient care. However, ethical considerations must be taken into account to ensure its responsible and beneficial application.
	Masters (2023) [57]	Addresses concerns about the improper utilization of ChatGPT by students and academics, emphasizing the importance of addressing this issue to prevent misuse and mitigate potential negative consequences.

The research papers in this section reveal a range of findings regarding the integration of GAI tools in medical education and practice. Notably, ChatGPT and GPT-4 showcase potential benefits, including personalized learning, improved comprehension of medical concepts, and effective translation of radiology reports. However, concerns about data privacy, potential bias, and impacts on critical thinking are brought to light. In clinical education, these tools exhibit promise in enhancing decision-making and supporting surgical training, although occasional random or oversimplified responses are identified as areas of caution. Specialized medical fields, such as neurosurgery and parasitology, show potential for GAI integration, but validation and context-specific refinement are crucial. The application of GAI in health education for individuals with Down syndrome highlights its significance in cognitive development and social interaction. Moreover, GAI's role in pandemic response and shaping future medicine proves noteworthy, with AI models contributing to mitigating COVID-19 spread and suggesting new research directions. Ethical considerations emerge as a common theme, emphasizing responsible usage and preventing misuse of GAI tools in healthcare.

A critical perspective reveals that these papers sometimes tend to overlook significant concerns such as data privacy, biases, and the risk of undermining critical thinking. Consequently, future work should prioritize establishing ethical guidelines and governance frameworks to address data privacy, bias, and transparency. Rigorous validation processes are essential to ensure the accuracy and reliability of GAI-generated information. Continuous research and development efforts are needed to enhance GAI capabilities for more complex medical scenarios. Interdisciplinary collaboration can promote responsible and effective integration of GAI in medical practices and education.

3.5. Application of GAI in Nursing Education

In this section, we explore a few research papers that delve into the implications of ChatGPT in nursing education (Table 5).

Table 5. The application of GAI in nursing education.

Publications	Authors	Focus
Potential and implications of ChatGPT	Choi et al.'s (2023) [58]	Examines the effects of ChatGPT on nurse education, highlighting the importance of striking a balance between AI integration and human interaction to ensure comprehensive learning experiences.
	Irwin et al. (2023) [59]	Discusses the broader implications of ChatGPT in nursing and midwifery practice, stressing the potential benefits it brings alongside the challenges in implementing this transformative technology.
Ethical considerations of using ChatGPT in nursing education	Abdulai and Hung (2023) [60]	Delves into the ethical considerations of using ChatGPT in nursing education, emphasizing the significance of aligning AI integration with nursing's core values to maintain the profession's unique identity.
Fostering critical thinking skills among nursing students.	Seney et al. (2023) [61]	Sheds light on ChatGPT's role in enhancing clinical judgment in nursing education and proposes a thoughtful teaching strategy to address its limitations in generating NCLEX-style questions.
	Ahmed (2023) [62]	Takes a closer look at ChatGPT's transformative impact on personalized learning and clinical decision-making for nurses while addressing the ethical considerations surrounding its use.
	Lo (2023) [63]	Introduces the CLEAR Framework, a valuable tool for equipping nursing students with critical thinking skills, enabling them to engage effectively with AI language models and promoting responsible AI use in nursing education.

The research papers in this section focus on the potential benefits of integrating GAI into nursing education. These studies explore how AI can enhance patient care, information literacy, and critical thinking skills among nursing students. The papers emphasize the importance of balancing AI integration with human interaction for comprehensive learning experiences and address ethical considerations. They delve into topics such as AI's impact on clinical judgment, broader implications in nursing practice, personalized learning, and clinical decision-making. Frameworks are introduced to equip nursing students with critical thinking skills for responsible AI engagement.

Future work in this domain is needed to address critical issues such as data privacy, biases, and the potential threat of compromising human-centered nursing practices. This can be achieved by establishing robust ethical guidelines and governance frameworks for the integration of AI in nursing education. Additionally, rigorous validation processes are essential to ensure the accuracy and reliability of AI-generated content.

3.6. Application of GAI in Communication

In this section, we explore a collection of research papers that delve into the applications and implications of GAI in communication education. These studies shed light on how GAI models are revolutionizing research, education, journalism, and social media in the context of communication, as shown in Table 6.

Table 6. The application of GAI in communication.

Publications	Authors	Focus
GAI models are revolutionizing in the context of communication.	Wang (2022) [64]	Focuses on machine translation, specifically examining the use of generative adversarial network-based neural machine translation models to improve Chinese–English translation accuracy and naturalness.
	Santandreu-Calonge et al. (2023) [65]	Investigates the advantages of natural language processing tools, like ChatGPT, in addressing challenges such as jargon and coordination issues. The authors emphasize the importance of combining AI technologies with human expertise to achieve effective communication and enhanced patient outcomes.
GAI models are revolutionizing journalism	Pavlik (2023) [66]	Critically analyzes ChatGPT and discusses its implications. The analysis highlights the transformative capacity of GAI in shaping the future of journalism and media, highlighting both the opportunities and challenges that arise from the integration of AI in these fields.
	Feng et al. (2023) [67]	Conducts a data-driven analysis using Twitter and Reddit. Their research uncovers the profound influence of ChatGPT on streaming media and highlights the potential of integrating ChatGPT with visual generative models.

These studies showcase the effectiveness of GAI in improving machine translation accuracy and its transformative potential in journalism and media education. Furthermore, the impact of AI on streaming media is explored, providing an understanding of GAI's role in communication education.

The small number of papers provide a restricted view of the challenges and ethical issues linked to biases and human interaction. These papers predominantly accentuate the positive impacts of GAI. This highlights the need for further research to gain a holistic understanding of GAI's impacts on communication and strike a balance between technological advancements and maintaining authentic human engagement. Future research could focus on developing and assessing AI-powered tools for communication skills, as well as examining ethical implications and potential biases in GAI model implementation in communication and journalism.

3.7. Application of GAI in Academia

This section examines a collection of research papers focusing on GAI's impact on research and its potential applications in different fields, as summarized in Table 7 below.

This section highlights the transformative impact of GAI in academia, showcasing its potential benefits for higher education, research, and library services. The discussion covers the advantages of AI technologies, natural language processing techniques, and large language models in various domains. However, it also stresses the need to address ethical concerns, algorithmic biases, and maintain academic integrity. Strategies to ensure integrity and accuracy are explored to address these challenges. Additionally, the section sheds light on the implications of GAI in library science.

Moving forward, future research should take a critical and interdisciplinary approach to GAI integration in academia, considering a wide array of potential benefits and risks. It is imperative to delve deeper into the ethical implications, potential biases, and algorithmic transparency associated with GAI tools. Moreover, a more in-depth investigation into the long-term effects of GAI on academic integrity, learning experiences, and research methodologies is warranted.

Table 7. The application of GAI in academia.

Publications	Authors	Focus
GAI and ethical considerations in academia	Alqahtani et al. (2023) [68]	Highlights the transformative impact of (GAI) models, such as GPT-4 and BARD, in higher education and research. They explore the potential benefits of these technologies, including text generation, data analysis, literature review, and educational support. However, the authors also emphasize the importance of addressing ethical concerns and algorithmic biases to ensure responsible and effective utilization of GAI in academia.
	Cotton et al. (2023) [69]	Sheds light on the benefits and challenges associated with the use of ChatGPT in higher education. The authors emphasize the advantages of increased student engagement and accessibility while addressing concerns related to academic dishonesty. Strategies to detect and prevent cheating are explored, highlighting the need for a proactive and ethical approach to integrating AI tools.
	Currie (2023) [70]	Focuses on the impact of ChatGPT on academic integrity in nuclear medicine and radiology, acknowledging its potential while highlighting limitations, such as errors and fabrications. Emphasizes the ethical and responsible integration of ChatGPT.
	Dergaa et al. (2023) [71]	Examines the prospects and potential threats of ChatGPT in academic writing, emphasizing the need to redefine norms and adjust information expectations for ethical utilization.
	Perkins (2023) [72]	Explores the implications of AI tools like ChatGPT on academic integrity during formal assessments, addressing concerns about undetectable original content generation. The importance of transparent student use and updates to academic integrity policies in higher education institutions are emphasized throughout these papers.
GAI in library science	Inamdar (2023) [73]	Conducts a comprehensive investigation into the impact of artificial intelligence text generators (AITGs) on library services, resources, and staff roles, highlighting their potential to enhance various aspects of library operations and accessibility.
	Frederick (2023) [74]	Focuses on information environments and explores the role of ChatGPT in libraries, analyzing current discussions and historical developments. The paper uncovers the potential risks and benefits associated with ChatGPT, providing valuable guidance for librarians to navigate the implications of this technology and emphasizing the importance of librarians acquiring knowledge about AI to adapt to changing information landscapes.

3.8. Application of GAI in General Education

In this section, we delve into the general applications of GAI in education. While earlier sections have explored the use of GAI in specific domains, such as medical education, higher education, engineering education, and more, this section broadens the scope to encompass the wider applications of GAI in education. To provide a structured analysis of this topic, the content is divided into several subsections. This section, therefore, covers a review of a wide range of applications and considerations regarding the integration of GAI in education, as discussed in Table 8 below:

Table 8. Application of GAI in general education.

GAI Application	Authors	Focus
Technology Adoption and Integration in Education	Al Ghatrifi et al. (2023) [75], Geerling (2023) [76], Skavronskaya et al. (2023) [77], Elnozahy and El Khayat (2023) [78], Aydın and Ayhan Erdem (2022) [79], Cooper (2023) [80], Su and Yang (2023) [81], Murugesan and Cherukuri (2023) [82], Hsu and Ching (2023) [83], Mørch and Andersen (2023) [84], Gentile et al. (2023) [85], Jangjarat et al. (2023) [86], Tlili et al. (2023) [87], Lodge et al. (2023) [88], Chen et al. (2020) [89], Luo et al. (2023) [90], Kasneci et al. (2023) [91], Golovianko et al. (2023) [92], Ye and Bors (2022) [93].	• AI, blockchain, and big data analysis applicability in accounting education systems. • ChatGPT and AI as tools for sophisticated evaluation methods. • Ethical integration of AI in tourism education. • AI decision-making support across various languages and contexts. • AI technologies in different linguistic and cultural settings. • Benefits and drawbacks of integrating LLMs, like ChatGPT, in education. • Educative AI framework that highlights personalized learning experiences and efficient feedback. • Responsible GAI implementation while optimizing the advantages of GAI in an ethical manner. • Uses, benefits, and concerns of GAI in education. • Human-centered AI approach in education to balance technology integration with pedagogical principles and learner needs. • Manifesto to guide educators through the AI transformation; focuses on proactive measures to integrate AI into teaching practices. • Public's acceptance of ChatGPT and the importance of considering human interaction factors. • Importance of responsible integration and the need for further research to ensure safe and ethical use of AI in education. • Impact of GAI on tertiary education. • Review of artificial intelligence in education. • ChatGPT's roles and challenges in early childhood education. • Opportunities and challenges of large language models, including ChatGPT, in education. • Pi-Mind agent, a cognitive digital clone designed to emulate human decision-making processes. • Teacher–student network framework that combines GAN and Variational Autoencoder (VAE) to enable lifelong learning.
Classroom Design and Learning Environment	Karadag et al. (2022) [94], Vartiainen and Tedre (2023) [95], Zhai et al. (2023) [96], Sarr et al. (2020) [97].	• Education AI model for generating adaptable and functional classroom layouts. • Using text-to-image GAI in craft education. • MACHE-Bot, an AI dialogue system, enhances the learning experience through humor and empathy. • Djehuty, an AI-powered educational gamified environment that addresses literacy challenges in resource-constrained settings.
Assessment Practices and Academic Integrity	Geerling (2023) [76], Bauer et al. (2023) [98], Farrokhnia et al. (2023) [99], Howell and Potgieter (2023) [100], Huang and Khan (2023) [101], Ibrahim et al. (2023) [102], Sung et al. (2019) [103], Lo (2023) [104], Ajevski et al. (2023) [105], Johinke et al. (2023) [106].	• Questioning traditional assessment methods and highlighting AI's potential in education. • A comprehensive framework to enhance peer-feedback processes. • ChatGPT's strengths in generating credible responses and personalized learning experiences and limitations in comprehension and potential biases. • Concerns about the authenticity and credibility of AI-generated student work. • AI and machine learning techniques for assessing complex competencies in education. • Evaluating ChatGPT's efficacy in university-level courses. • Short answer grading using transformer-based pre-training. • Updating assessment methods, implementing policies, and educating instructors and students on ChatGPT's implications. • Adapted assessment strategies and responsible integration of AI. • Importance of upholding educational values while incorporating AI-powered writing generators, like ChatGPT.

Table 8. *Cont.*

GAI Application	Authors	Focus
Language Instruction	Kovačević (2023) [107], Zhao and Song (2022) [108], Tang and Deng (2022) [109], O'Leary (2023) [110], Yan (2023) [111].	• The potential of ChatGPT in English for Specific Purpose (ESP) teaching and learning and how it enhances lesson planning, execution, and assessment of written assignments. • AI-assisted feedback system for English language classes. • AI-based expert system to enhance students' interest, learning strategies, and English language abilities. • Comparative analysis of language chatbots. • ChatGPT's potential to enhance writing efficiency and workflow in Language 2 (L2) writing classrooms, concerns about academic integrity, and the need for policies and pedagogical guidance for responsible usage.
Ethics and Responsible GAI Use	Kooli (2023) [13], Bearman and Ajjawi (2023) [112], Fernandez (2023) [113], Bahrini et al. (2023) [3], Ray (2023) [114], Singh and Singh (2023) [115], Danry et al. (2022) [116], Ali et al. (2021) [117].	• The need for adaptation, heightened awareness, and thoughtful ethical considerations to harness AI's positive impact. • A pedagogical approach to equip learners with skills to navigate AI systems. • Understanding societal boundaries and ethical standards to foster critical engagement. • Opportunities, threats, and limitations of ChatGPT, and the importance of human expertise. • Balance between AI-assisted innovation and human expertise. • The need for multidisciplinary research to understand and harness its potential. • Positive applications and ethical implications of AI-generated characters, including design considerations and potential impact in entertainment, education, and therapy. • Understanding GAI techniques and assessing the societal implications of misinformation through a curriculum for middle school students.
Subject-Specific Education and Skills Development	Adams et al. (2023) [118], Bearman and Ajjawi (2023) [112], Emenike and Emenike (2023) [119], Agathokleous et al. (2023) [120], Kooli (2023) [13], Hallsworth et al. (2023) [121], Megahed et al. (2023) [122], Ivanov and Soliman (2023) [123], Fergus et al. (2023) [124], Siche and Siche (2023) [125], Uddin et al. (2023) [126], Wolfert et al. (2022) [127].	• Enriching curricula with sustainable development goals (SDGs) through improved keyword analysis. • Equipping learners with skills to navigate AI systems and understand ethical standards associated with AI in a relational context. • Implications of AI programs, like ChatGPT, in chemistry education. • Potential advantages of ChatGPT and similar AI technologies in biology and environmental science. • Effective regulations and ethical values to safeguard educational integrity. • The role of theory-based approaches in biology, the importance of human engagement in scientific innovation, and the value of non-mainstream research. • Application of ChatGPT in the domain of statistical process control (SPC). • Implications of ChatGPT in tourism education and research. • Application of ChatGPT in addressing chemistry assessment queries. • Potential applications of ChatGPT in the agricultural and livestock industry. • Effectiveness of ChatGPT in enhancing hazard recognition capabilities among construction students. • Integrating Building Information Models (BIM) with diverse information systems in construction projects by introducing the novel concept of Open Design Learning (ODL).

Table 8. *Cont.*

GAI Application	Authors	Focus
Learning Outcomes and Student Engagement	Wu and Yu (2023) [128], Muñoz et al. (2023) [129], Crawford et al. (2023) [130], Shoufan (2023) [131], Gardner and Giordano (2023) [132], Hemachandran et al. (2022) [133], Limo et al. (2023) [134], Alnaqbi and Fouda (2023) [135], Jeon and Lee (2023) [136], Dijkstra et al.(2023) [137], Pataranutaporn et al. (2022) [138].	• Human-like avatars, gamification elements, and emotional intelligence in chatbot design to further enhance learning outcomes. • Enhanced learning outcomes through the integration of ChatGPT into education. • ChatGPT for character development and authentic assessments in higher education. • Senior computer engineering students' perceptions of using ChatGPT in learning activities. • The importance of communication and teamwork skills is highlighted by the value of oral exams in physical chemistry classrooms. • The need for improved communication and information exchange within the classroom. • Influence of factors such as students' perception, satisfaction, and engagement on the utilization of ChatGPT for personalized tutoring. • Utilization of ChatGPT and social media for collecting real-time student feedback on teaching styles in higher education. • Collaborative nature of teacher–AI interaction and the importance of teachers' pedagogical expertise. • EduQuiz, an all-encompassing quiz generator that utilizes a fine-tuned GPT-3 model. • Influence of learning from an AI-generated virtual instructor resembling an admired individual.
Challenges and Considerations in AI Integration	Shu et al. (2021) [139], Wardat et al. (2023) [140], Halaweh (2023) [141], Lim et al. (2023) [142], Hwang and Chen (2023) [143], Yu and Guo (2023) [144], Dwivedi et al. (2023) [145].	• Vulnerability of STEM learners to AI-generated misinformation. • Perspectives of students and educators on using AI. • Analyzed the concerns expressed by educators regarding the integration of ChatGPT in educational settings. • A comprehensive framework for the future coexistence of GAI and education. • GAI applications, challenges, and future research directions. • GAI and ChatGPT's current status, issues, and prospects in educational settings. • Opportunities, challenges, and implications of generative conversational AI for research, practice, and policy.
Pedagogical Approaches and Teaching Methods	Javed et al. (2022) [146], Bauer et al. (2023) [98], Karaali (2023) [147], Lawrie (2023) [148], Robertson and Muirhead (2022) [149], Álvarez-Álvarez and Falcon (2023) [150].	• Proposed the BAG model (Build, Assess, and Govern) as a teaching paradigm. • A comprehensive framework to enhance peer feedback processes. • Challenges of teaching quantitative literacy and reasoning alongside language-based AI systems, like ChatGPT. • Teachers' adaptability and creativity in incorporating AI tools into their instructional approaches. • Gaps in curriculum and educational policies, emphasizing the need for interdisciplinary approaches. • Students' preferences regarding university teaching methods.

The remaining part of this section synthesizes and discusses the applications of GAI in general education across each subcategory of papers outlined in Table 8.

Technology Adoption and Integration in Education—The findings demonstrate improved learning outcomes, enhanced student engagement, and the potential for personalized learning experiences. However, ethical considerations, responsible usage, and addressing limitations are crucial. Future work necessitates a more critical perspective in developing nuanced guidelines and policies, cautiously navigating interdisciplinary collaboration and fostering genuinely informed critical thinking skills. This is imperative to ensure that AI integration in education addresses complex challenges effectively and ethically. Additionally, an in-depth exploration of the lasting repercussions of AI technologies on education is crucial, surpassing superficial assessments and delving into their potential to reshape fundamental educational dynamics.

Classroom Design and Learning Environment—The papers in this category highlight the significance of incorporating technology and GAI in creating effective learning environments. The research demonstrates the potential of AI models to automate classroom design, optimize learning spaces, and enhance the learning experience through interactive dialogue systems and gamified environments. However, it is crucial to acknowledge the potential pitfalls, like algorithmic biases and copyright concerns. Moving forward, there is a pressing need to prioritize the promotion of AI interactions that are inclusive and culturally sensitive. Moreover, the development of AI-powered solutions tailored to resource-limited educational contexts should be a key focus of future endeavors.

Assessment Practices and Academic Integrity—The research highlights the superior performance of AI chatbots, such as ChatGPT, in subject-specific exams and the potential of natural language processing techniques to enhance peer feedback processes. However, concerns regarding authenticity, credibility, systematic errors, and the risk of plagiarism are raised, underscoring the need for caution and the responsible integration of AI tools in educational assessments. While the papers touch on the importance of updating assessment strategies, policy implementation, instructor training, and student awareness about AI's impact on evaluations, a more in-depth analysis of these measures' efficacy is warranted. To advance, it is crucial to fine-tune AI models for assessment precision, tackle biases and accuracy concerns, and establish robust guidelines and ethical standards to govern AI's role in educational assessments.

Language Instruction—The publications demonstrate how integrating AI chatbots into language instruction can offer personalized support, optimize learning experiences, and address diverse language proficiencies. However, a more critical viewpoint uncovers the apprehensions surrounding the preservation of academic integrity and equitable learning experiences. While the potential gains are evident, the unexplored territory of AI's potential biases and fairness concerns necessitates not only policy frameworks but also a deep reevaluation of instructional strategies.

Ethics and Responsible GAI Use—These papers highlight the importance of ethical considerations and the responsible use of AI systems, particularly chatbots like ChatGPT, in education and research. These works shed light on the merits, shortcomings, obstacles, and debates surrounding such technologies. Looking ahead, it is imperative that forthcoming efforts concentrate on interdisciplinary investigations to confront ethical difficulties, integrate conscientious AI education into academic curricula, and promote collaborative endeavors aimed at formulating comprehensive regulations and principled directives.

Subject-Specific Education and Skills Development—The reviewed research papers shed light on the potential applications of GAI in specific educational domains, including sustainability education, digital skills development, chemistry education, biology and environmental science, and construction education. The findings indicate that GAI technologies, such as ChatGPT, have the ability to simplify complex tasks, expedite education and research processes, enhance hazard recognition capabilities, promote safety management practices, and facilitate interdisciplinary approaches. However, the papers also acknowledge the limitations and potential risks associated with the utilization of GAI, such as concerns about data privacy, biases, and the need for human engagement. Moving forward, future work should focus on the seamless integration of sustainability education into AI learning frameworks, a dedicated commitment to addressing issues of equity and accessibility to ensure that the benefits of GAI are universally accessible, and a diligent pursuit of strategies to maximize the advantages that GAI offers while concurrently taking proactive steps to mitigate the inherent risks that come with its implementation.

Learning Outcomes and Student Engagement—The papers highlight the positive impact of GAI, such as AI chatbots and virtual instructors, on students' learning outcomes and engagement. The findings demonstrate improvements in learning outcomes, motivation, and engagement when utilizing AI chatbots, like ChatGPT. However, a more discerning analysis beckons for heightened scrutiny in various aspects. This includes the imperative refinement of AI chatbot design, probing into alternative assessment strategies, enhancing the interaction between AI and educators, and undertaking an in-depth exploration into the long-term impact of AI on student motivation and engagement.

Challenges and Considerations in AI Integration—The related papers shed light on the obstacles and considerations associated with integrating GAI, such as ChatGPT, into education. These studies emphasize the importance of addressing concerns related to AI-generated misinformation, ensuring the secure integration of AI in education, and promoting responsible implementation practices. The findings underscore the need for raising learners' awareness about the potential influence of AI on their learning and further research to ensure the conscientious and secure integration of AI technologies, like ChatGPT, in educational settings. The future trajectory in this field involves developing comprehensive strategies and guidelines to address challenges while conducting further research on AI integration's impact on student learning outcomes. This avenue of exploration aims to formulate effective approaches to counter challenges and rigorously investigate how AI integration influences students' academic achievements.

Pedagogical Approaches and Teaching Methods—The papers shed light on innovative instructional practices and frameworks that integrate AI technologies into education. The findings offer insights into teaching ethics in AI courses, enhancing peer feedback processes through natural language processing techniques, addressing challenges in teaching quantitative literacy alongside language-based AI systems, and showcasing the integration of AI to support student learning. The papers also emphasize the importance of interdisciplinary collaboration. Future endeavors could dedicate their focus to meticulously refining and substantially expanding the currently presented pedagogical methodologies. Delving deeper, a crucial area of investigation could encompass understanding the enduring effects of AI integration on the overall learning achievements of students, shedding light on potential shifts in learning dynamics, cognitive development, and critical thinking skills over the long term. Furthermore, it becomes imperative to navigate the intricate field of ethical considerations, delving into the multifaceted implications of AI technologies within educational contexts.

To sum-up the findings in this section, Figure 3 provides a comprehensive visualization and categorization of the diverse applications of GAI in the field of education. This spider graph employs a two-tier structure to present the information effectively. The initial tier highlights the primary themes that have been central to a majority of applications. The proportional sizing of the circles reflects the prevalence of these themes, representing the number of applications and associated publications in each domain. Furthermore, within each thematic category, a secondary layer delineates specific keywords, encapsulating the nuances of each area. The size of these sub-circles corresponds to the prominence of these keywords, which are characterized by their density and the number of related research publications. This elaborate spider graph succinctly encapsulates the breadth and depth of GAI's influence in education, providing an illuminating snapshot of the ongoing research endeavors and practical implementations in this field.

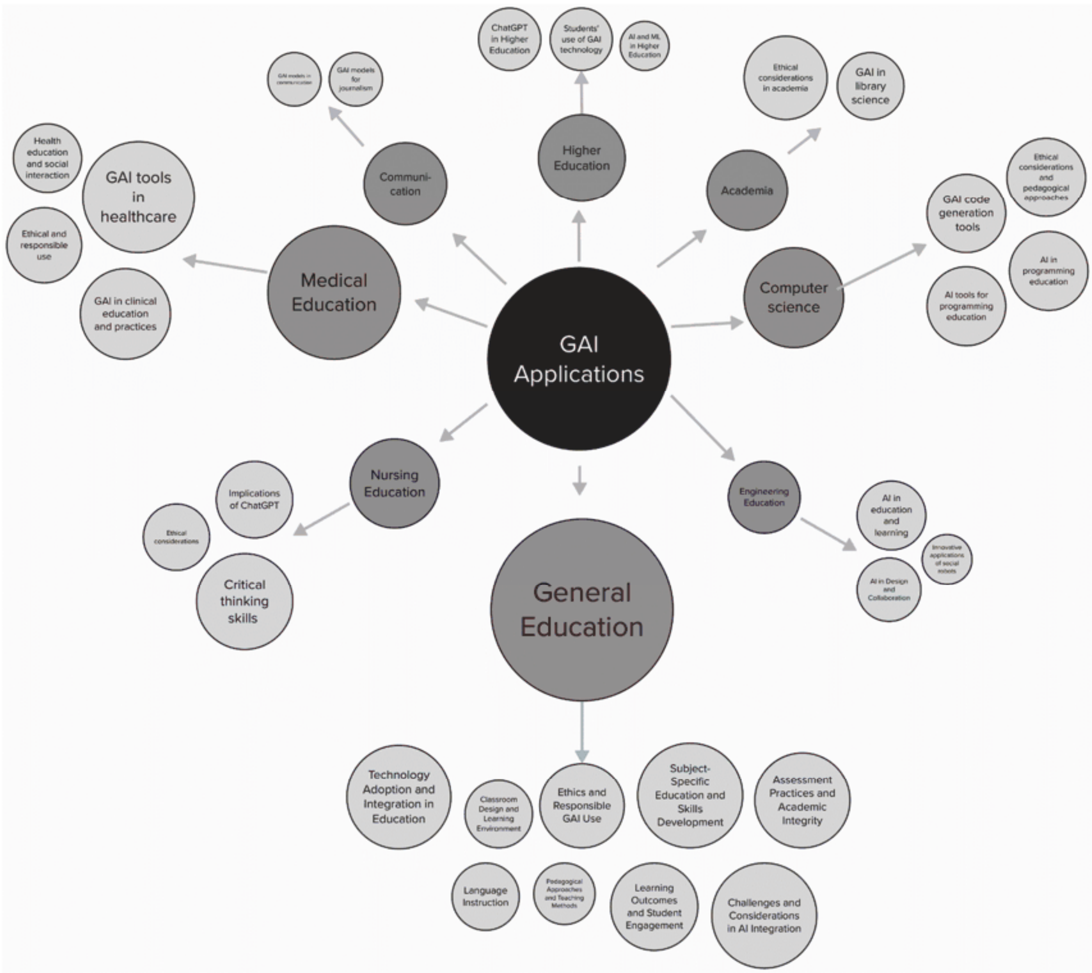


Figure 3. Summary of GAI research and applications in education.

4. Bibliometric Analysis

In this section, we embark on a comprehensive bibliometric analysis that delves into the field of GAI within the context of education. By examining a vast collection of existing articles and publications retrieved from the Scopus database, we seek to provide a deeper understanding of the current state of research and innovation in this evolving field. Through systematic literature screening and data analysis, we aim to uncover valuable insights, trends, and implications surrounding GAI's integration and impact on teaching and learning in educational settings. This exploration will shed light, for example, on the interdisciplinary nature of GAI research, key AI tools employed, the geographic distribution of contributions, and the diverse fields within education where GAI finds its applications. As we navigate through the findings and visualizations, we uncover the growth trajectory of GAI research, observe patterns of co-occurrence between terms, and ascertain the prominent applications of GAI tools, like ChatGPT, in educational settings.

To provide an in-depth analysis of the current state of GAI in education, we utilized VOSviewer to generate five types of visualization maps. These maps are designed to provide a comprehensive analysis of the topic. Each map consists of circles that represent different items such as publications, researchers, or terms. The size of the circle and the font within it indicate the level of activity associated with that item. A larger circle and a bigger font size indicate a higher level of activity, whereas a smaller circle and a smaller font size suggest less activity. The distance between any two terms in the diagram reflects the degree of association between them. A shorter distance signifies a stronger correlation, whereas a longer distance indicates a weaker correlation [151].

4.1. Co-Occurrence Map Based on Text Data

The relevant and frequently occurring terms were identified by analyzing the text data of the 217 selected publications. The analysis of text data looks to extract relevant terms from the titles and abstracts of the selected articles and constructs a network of co-occurrence links between them [151]. This analysis enables us to detect any emerging developments and identify impactful terms within the field of GAI in education. A total of 4408 terms were generated, out of which 102 terms meet the minimum threshold of 10 occurrences. To further narrow down the selection of terms, VOSviewer calculated a relevance score for each term, and the top 60% most relevant terms were selected, resulting in the 61 terms that were displayed in the network seen in Figure 4. Terms with a high relevance score represent more specific topics covered by the text data, while terms with a low relevance score tend to be of a general nature [151]. A thesaurus was created and uploaded to VOSviewer for the standardization of keywords and the elimination of duplicate terms [152].

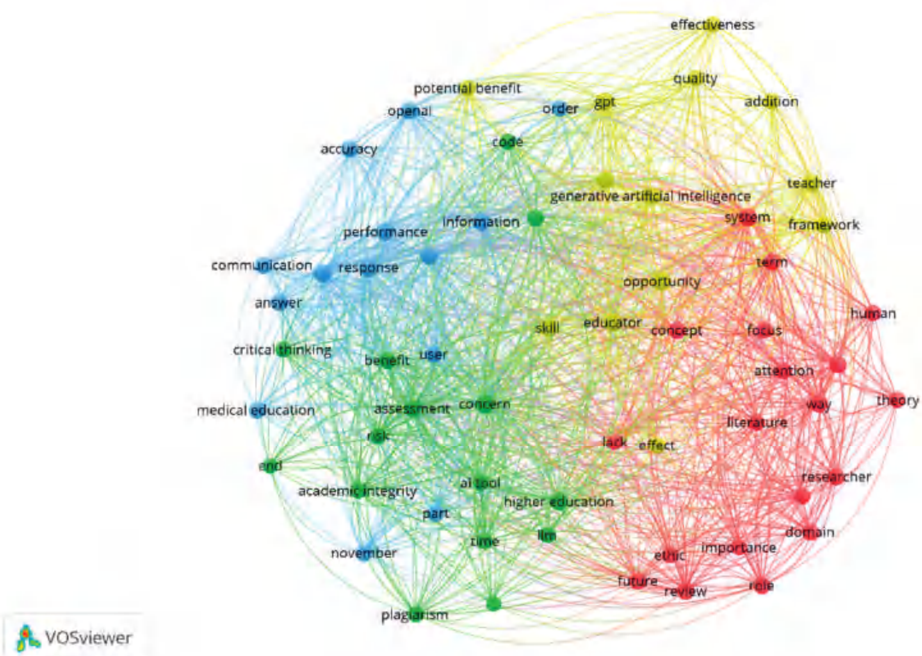


Figure 4. Co-occurrence map of text data.

The findings in Figure 4 demonstrate the diverse range of research and exploration within the field of GAI in education. The interconnected web of key terms suggests that the research on GAI has encompassed the areas of assessment, teaching, research, higher education, academic integrity, and ethics. The findings also highlight GAI's revolutionary potential with the appearance of terms such as "opportunity" and "effectiveness". Additionally, the terms that are directly linked with GAI are shown in Figure 5.

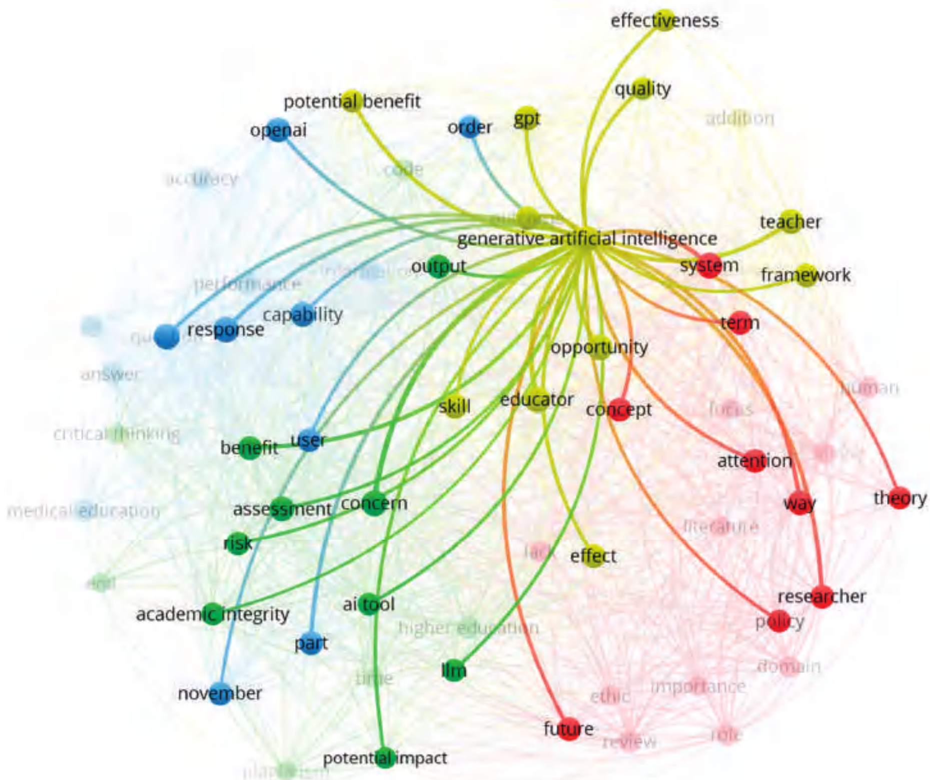


Figure 5. Terms directly connected with "Generative Artificial Intelligence".

The findings in Figure 5 indicate a strong relationship between GAI and "Generative Pre-training Transformers" (GPT), which are a type of large language model (LLM). GPTs, such as ChatGPT, have emerged as foundational technologies in GAI and have demonstrated their capabilities in various applications. The integration of GPTs in educational contexts, particularly in the development of AI chatbots, has the potential to enhance learning experiences through personalized interactions, intelligent feedback, and adaptive learning pathways. This raises questions about the impact of GAI on student engagement, motivation, and learning outcomes. Additionally, the association between GAI and terms such as opportunity, question, response, and assessment highlight the diverse applications of GAI in educational contexts. GAI has the potential to support various aspects of teaching and learning, including generating questions, providing responses, and assessing student performance. The use of GAI in assessments raises concerns about the fairness and reliability of automated grading systems. It is crucial to examine the validity and equity of GAI-based assessment methods to ensure that they align with established educational standards and principles of fairness.

The ten most frequently occurring terms and their number of occurrences can be seen listed below in Table 9. Also listed in the table is the relevance score for each term. The relevance score for each term is calculated by VOSviewer. For each term, VOSviewer determines the distribution of (second order) co-occurrences over all terms and then compares it to the overall distribution of co-occurrences over terms. The larger the difference between the two distributions (measured using the Kullback–Leibler distance), the higher the relevance of a term [153]. Table 9 provides an insightful analysis of the key terms frequently encountered in the context of GAI in education. These terms shed light on the diverse aspects of GAI's integration into educational practices. The term "System" suggests the broader framework within which GAI operates, emphasizing the need for systematic implementations. "Response" indicates the generation of AI-driven outputs, while "Question" signifies GAI's potential to create queries and foster interactive learning environments. "Concern" highlights the importance of addressing ethical and practical considerations when employing GAI in education. "Opportunity" underscores the positive prospects that GAI offers in enhancing educational experiences and outcomes. The prominence of "Generative Pre-training Transformers (GPT)" demonstrates the pivotal role of advanced language models, like ChatGPT, in driving GAI applications. Terms like "Assessment" and "Capability" indicate GAI's relevance in evaluating student performance and its potential to augment educational capabilities. This analysis unveils the nuanced facets of GAI in education, encompassing both its capabilities and the essential aspects that demand attention for responsible implementation.

Table 9. Top 10 terms by occurrences.

Rank	Term	Occurrences	Relevance Score
1	System	39	0.8012
2	Response	38	1.4718
3	Concern	35	0.8416
4	Question	34	1.1378
5	Opportunity	30	0.8361
6	GPT	28	1.0504
7	Information	27	1.4886
8	Assessment	26	0.6512
9	Way	25	1.134
10	Capability	25	0.5872

4.2. Co-Occurrences Map Based on Keywords

The keywords that occur frequently were identified by analyzing the bibliographic data of the 217 selected publications. A total of 1204 keywords were found, out of which 50 were selected, where the minimum number of occurrences of those terms was limited to 6 (Figure 6). A thesaurus was also used in VOSviewer to standardize keywords and eliminate redundant duplicate terms. This analysis looks at all keywords, including both author keywords and index keywords. Author keywords are specific keywords chosen and provided by the authors of the articles themselves, while index keywords are keywords that have been assigned by indexers or databases to classify and categorize the articles for the purpose of information retrieval and indexing.

Some of the most common keywords identified include "Artificial intelligence", "ChatGPT", "Education", and "Human". A list of the ten most frequently used keywords, their occurrences, and their total link strength can be found in Table 10. The total link strength indicates the total strength of the co-occurrence links of a given keyword with other keywords [151].

Table 10 presents a comprehensive analysis of the most frequently occurring keywords in the domain of GAI in education, offering valuable insights into the prominent themes and research areas. The top ten keywords are indicative of the central focus and key aspects of GAI's integration into educational contexts. "Artificial intelligence" emerges as the most prevalent keyword, reflecting the overarching role of AI technologies in transforming educational practices. "ChatGPT" follows closely, showcasing the widespread recognition and utilization of this AI chatbot in educational applications. "Education" and "Human" hold significant positions, highlighting the significance of GAI in shaping learning experiences and considering human aspects in the educational process. "Education computing" signifies the intersection of computing technologies and education, while "Machine learning" underscores the pivotal role of ML algorithms in GAI applications. "Medical education" and "Natural language processing" reveal GAI's potential in healthcare education and language-based learning environments. Lastly, "Students" emphasizes the focus on learners and their experiences, reflecting the student-centric approach that GAI-driven educational interventions strive to achieve. These keywords depict the breadth and depth of GAI research in education, with an emphasis on AI-driven technologies, human-centric design, and transformative potential in diverse educational settings.

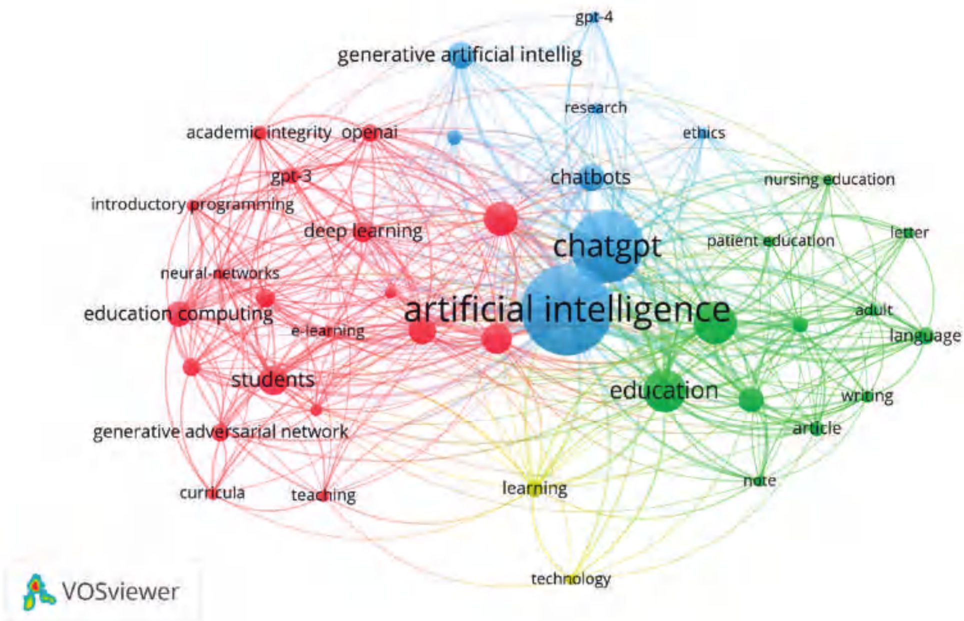


Figure 6. Co-occurrence map of all keywords.

Table 10. Top 10 keywords by occurrences.

Rank	Keyword	Occurrences	Total Link Strength
1	Artificial intelligence	111	97
2	ChatGPT	88	69
3	Education	41	37
4	Education computing	19	18
5	Human	41	41
6	Humans	26	26
7	Machine learning	21	18
8	Medical education	19	19
9	Natural language processing	19	18
10	Students	24	21

To further the analysis, two additional keyword maps were generated. Firstly, a map of the most common author keywords was generated. The keywords were limited to a minimum of five occurrences, which resulted in a total of twenty-five author keywords in the network visualization (Figure 7). Similarly, a map of the most common index keywords was created. The keywords were also limited to a minimum of five occurrences, which resulted in a total of thirty-nine index keywords in the map (Figure 8).

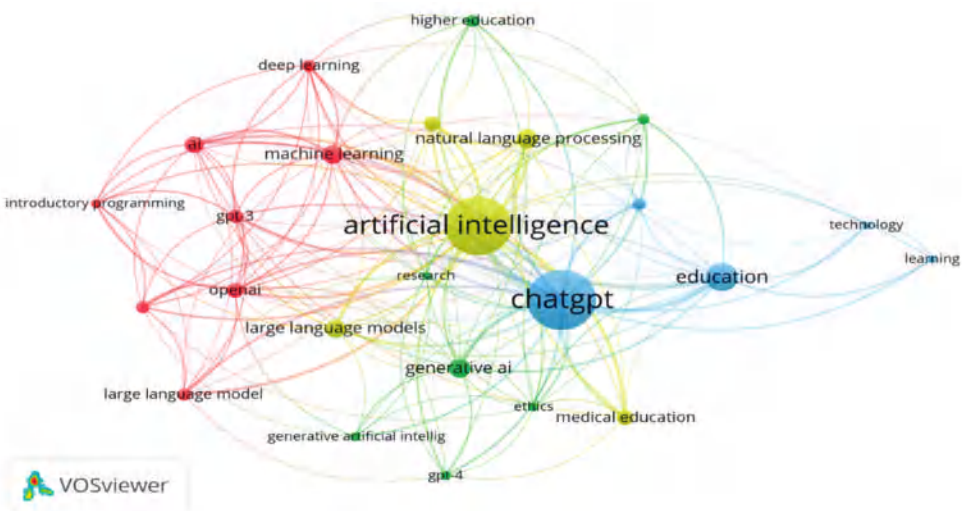


Figure 7. Co-occurrence map of author keywords.

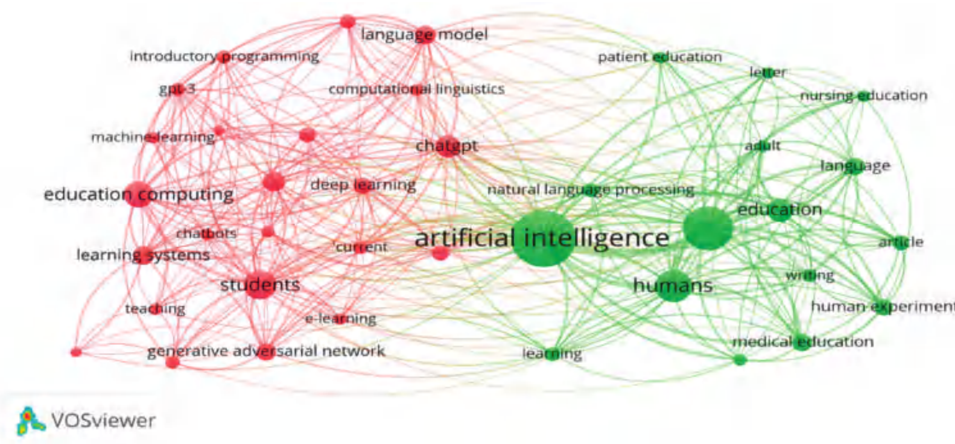


Figure 8. Co-occurrence map of index keywords.

The analysis of keywords reveals the central focus of GAI research in education, with terms such as artificial intelligence, ChatGPT, and education featuring prominently. This emphasizes the significance of AI technologies in transforming educational practices and the need for collaborations between AI experts and educators. The inclusion of medical education-related keywords suggests potential applications of GAI in healthcare education. GAI has the potential to support medical training, enhance diagnostic processes, and improve patient care.

4.3. Co-Occurrence Map Based on Country of Co-Authorship

Furthermore, the data analysis enabled us to examine the geographic distribution of these publications. The visualization map was generated for the countries of co-authorship by setting the minimum number of documents of a country to five. Of the 69 countries with publications, 20 met the threshold, and the results are depicted in Figure 9.

Figure 9. Country of co-authorships.

Most publications appeared from English-speaking countries (USA, Australia, and the UK), followed by India, Germany, and China. The ten countries with the strongest links on the map are listed in Table 11.

Table 11. Top 10 countries by link strength.

Rank	Country	Documents	Citations	Total Link Strength
1	United States	73	334	59
2	Australia	26	102	31
3	United Kingdom	15	89	29
4	New Zealand	8	67	27
5	India	15	98	20
6	Germany	11	75	19
7	Hong Kong	7	211	19
8	Italy	7	63	17
9	Saudi Arabia	6	40	17
10	Switzerland	5	45	16

Table 11 presents the top 10 countries based on their contributions to GAI research in education. The United States leads the list, indicating its central role in GAI-related studies. Australia and the United Kingdom follow closely, showcasing their active engagement in the field. New Zealand, India, Germany, Hong Kong, Italy, Saudi Arabia, and Switzerland also feature, reflecting their emerging roles in GAI research within the educational context.

The geographical distribution of publications demonstrates widespread global interest and collaboration in exploring the potential of GAI in education. Moreover, the distribution indicates the dominance of English-speaking countries such as the United States, Australia, and the United Kingdom. This aligns with the global prominence of these countries in educational research and technological advancements. However, it is worth noting the increasing contributions from countries like India, Germany, and China, suggesting a broader global interest in GAI in education. Collaboration and knowledge sharing across countries and regions can further propel advancements and foster cross-cultural perspectives in GAI research.

4.4. Co-Occurrence Map Based on Authorship

A further analysis looks at investigating the authorship relations across the multitude of authors of the publications. The network visualization map was limited to authors with a minimum of five citations. Out of the 682 authors, 144 met the threshold, and the largest set of connected authors included 23 authors. The resulting network visualization, depicted in Figure 10, presents a complex web of links connecting these 23 authors.

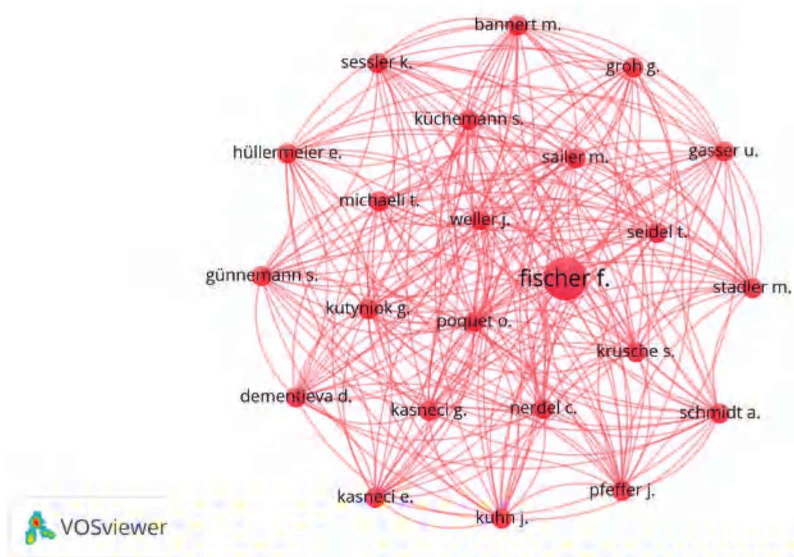


Figure 10. Co-occurrence map of authors.

The network visualization consists of a large network of links connecting the authors, which showcases the collaborative and interdisciplinary nature of the research being conducted within the field of GAI in education. Moreover, it demonstrates how researchers from diverse backgrounds and institutions come together to advance knowledge and innovation in this area. By highlighting these connections, the visualization fosters a better understanding of the knowledge-sharing dynamics and the collective effort behind the advancements in GAI in education. This insight can prove invaluable in identifying key contributors, potential research trends, and opportunities for further collaboration and development within the field.

The top 10 authors in terms of documents produced and citations can be seen in Tables 12 and 13, respectively. The total link strength in each table indicates the total strength of the co-occurrence links of a given author with other authors [151].

Table 12 depicts the key contributors of research in the domain of GAI in education. The relatively high number of documents, citations, and total link strength demonstrates the great influence and contribution of the authors.

Table 12. Top 10 authors by documents produced.

Rank	Author	Documents	Citations	Total Link Strength
1	Ali S.	4	23	13
2	Breazeal C.	4	23	13
3	Denny P.	4	30	12
4	Fischer F.	3	30	22
5	Becker B.A.	3	29	12
6	Dipaola D.	3	23	12
7	Finnie-Ansley J.	3	29	12
8	Luxton-Reilly A.	3	29	12
9	Prather J.	3	29	12
10	Cowling M.	3	9	3

Table 13. Top 10 authors by citations.

Rank	Author	Documents	Citations	Total Link Strength
1	Hwang G.-J.	2	162	3
2	Chen X.	1	162	3
3	Xie H.	1	162	3
4	Zou D.	1	162	3
5	Thorp H.H.	1	116	0
6	Stokel-Walker C.	1	51	0
7	Chartash D.	1	47	6
8	Chi L.	1	47	6
9	Gilson A.	1	47	6
10	Huang T.	1	47	6

Similarly, Table 13 identifies key contributors who have been influential through their publications within this domain. Notably, the top four authors, namely Hwang G.-J., Chen X., Xie H. and Zou D., all have an impressive citation count of 162. This exceptional citation count comes from their highly influential research paper titled "Application and theory gaps during the rise of Artificial Intelligence in Education", which played a vital role in exploring the applications of AI in education in 2020, contributing significantly to the early development of the field [89].

4.5. Data Analysis on Article Sources

To investigate the different sources of the 217 publications, a list of all unique sources was created and then ordered from the highest number of articles to the lowest. A total of 152 unique sources were found. The sources with the highest number of publications were "Computers and Education: Artificial Intelligence" and the "Journal of University Teaching and Learning Practice", with five publications each. A bar graph showcasing the top 10 sources by number of publications can be found in Figure 11.

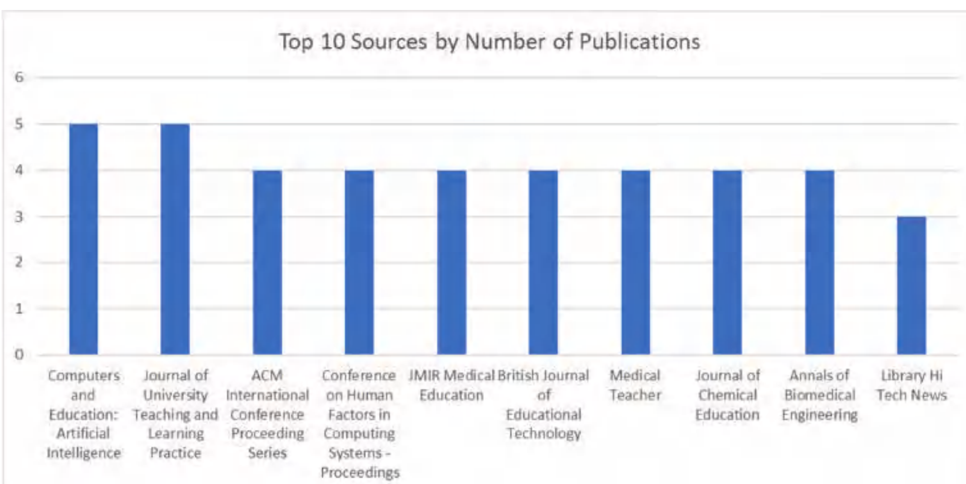


Figure 11. Bar graph of the top 10 sources by number of publications.

The analysis of article sources, therefore, presents a compelling picture of the interdisciplinary nature of GAI research in education, drawing insights from various fields such as computer science, education, and even the medical field. The convergence of expertise from diverse disciplines underscores the significance of collaborative efforts in addressing the complex challenges associated with GAI implementation in education.

4.6. Data Analysis on Document Type, Fields, GAI Tools Used, and Research Types

In this section, we classified the 217 publications based on their type, field of study, GAI tools used, and paper types. Figure 12 presents an analysis of the occurrences of various document types from Scopus. Notably, the most common document type is "Article", which appears 99 times, highlighting its preeminent role as the primary form of publication within the dataset. "Article" is followed by "Note", with 23 occurrences, and "Conference Paper" and "Review", with 20 and 18 occurrences, respectively. Moderately represented are "Letter" and "Editorial", each occurring 17 and 11 times, respectively. Conversely, "Short Survey" is the least common, appearing only once. The dominance of scholarly articles in the dataset aligns with the typical composition of academic databases like Scopus, emphasizing their pivotal role as reliable sources of scientific information.

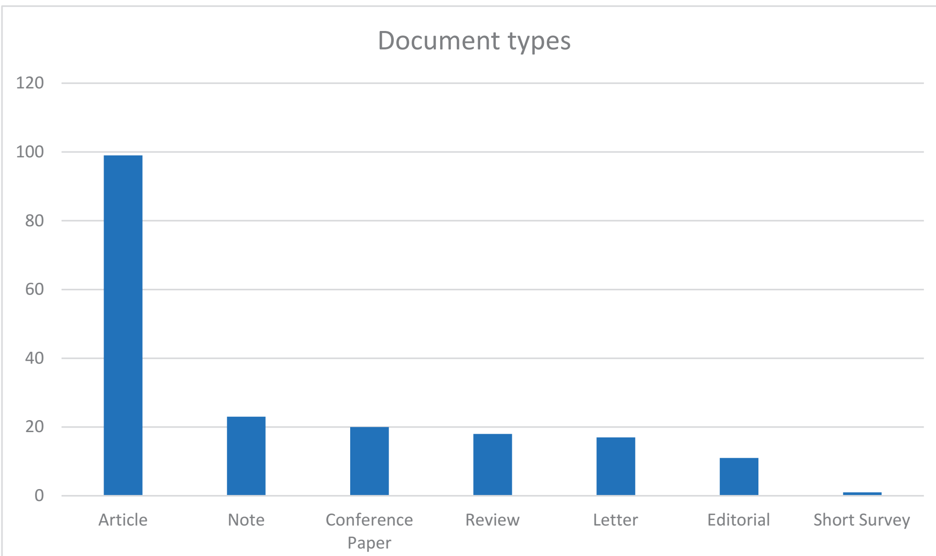


Figure 12. Occurrences of document types.

Additionally, the analysis delves further into the field, revealing 21 distinct disciplines within the dataset, with the top 10 most prevalent fields illustrated in Figure 13. It is crucial to acknowledge that papers falling into multiple categories, such as higher education and academia, were individually counted as separate occurrences.

As depicted in Figure 13, the majority of the papers encompass a broad scope of education and fall under the "general education" category, comprising 104 occurrences. Additionally, our analysis reveals the integration of GAI in various educational frameworks across different fields. Particularly noteworthy is the significant number of papers originating from the medical field, with 43 occurrences, indicating the growing adoption of GAI in both medical and educational contexts. Academia-related papers account for 21 occurrences, further highlighting the relevance of GAI in academic research. Moreover, we observed GAI's intersection with specific disciplines, such as computer science education, nursing education, and engineering education, with thirteen, nine, and nine occurrences, respectively. Lastly, it is worth mentioning the representation of GAI in other fields, including higher education and communication education, with eight and six occurrences, respectively, showcasing the diverse applications and potential of this technology in various educational domains.

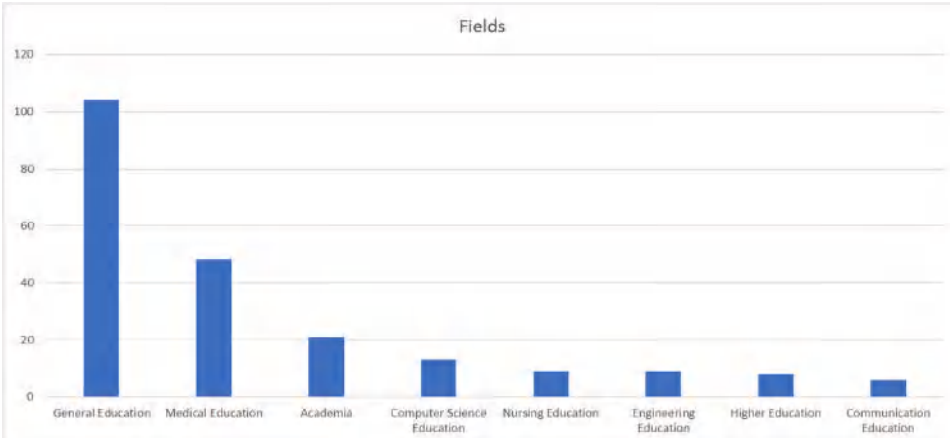


Figure 13. Occurrences of unique fields in education.

Moving forward, we conducted an analysis of the AI tools utilized in each paper, providing valuable insights into the tool landscape within the field of GAI in education (Figure 14). It is important to note that some papers utilize multiple AI tools, with each tool recorded as a distinct occurrence. Particularly noteworthy is the dominance of ChatGPT, which features in a significant number of papers, with a staggering 151 occurrences. This remarkable prevalence highlights the widespread recognition and adoption of ChatGPT within educational contexts, underlining its efficacy in generating coherent and contextually relevant responses for educational applications.

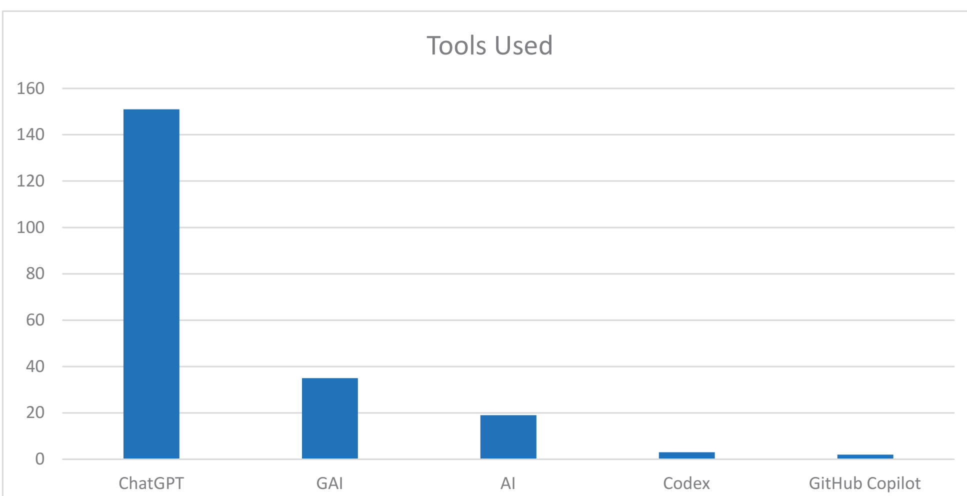


Figure 14. Occurrences of AI tools.

AI-powered Chatbots, such as BlenderBot and LaMDA, and AI text generators (AITG) appear in the dataset, albeit with lower frequencies [73,110]. Tools such as Codex and GitHub Copilot, which translate natural language to code, appear in a limited number of papers [5,154]. BERT, a method of pre-training language representations, and Conditional GAN, a type of GAN that involves the conditional generation of images by a generator model, appear a few times in the dataset. Other unique tools with a low number of occurrences include EDU-AI, a model for generating adaptable and functional classroom layouts [94], MACHE—bot, an AI dialogue system [96], and Djehuty (AI), an AI-powered educational gamified environment [97]. Furthermore, several papers do not specify a particular tool and instead use more general terms, like "GAI" and "AI", to describe the approaches they employed.

Then, we outlined the types of research in the context of GAI in education to gain valuable insights into the nature and focus of the scholarly work conducted in this field (Figure 15).

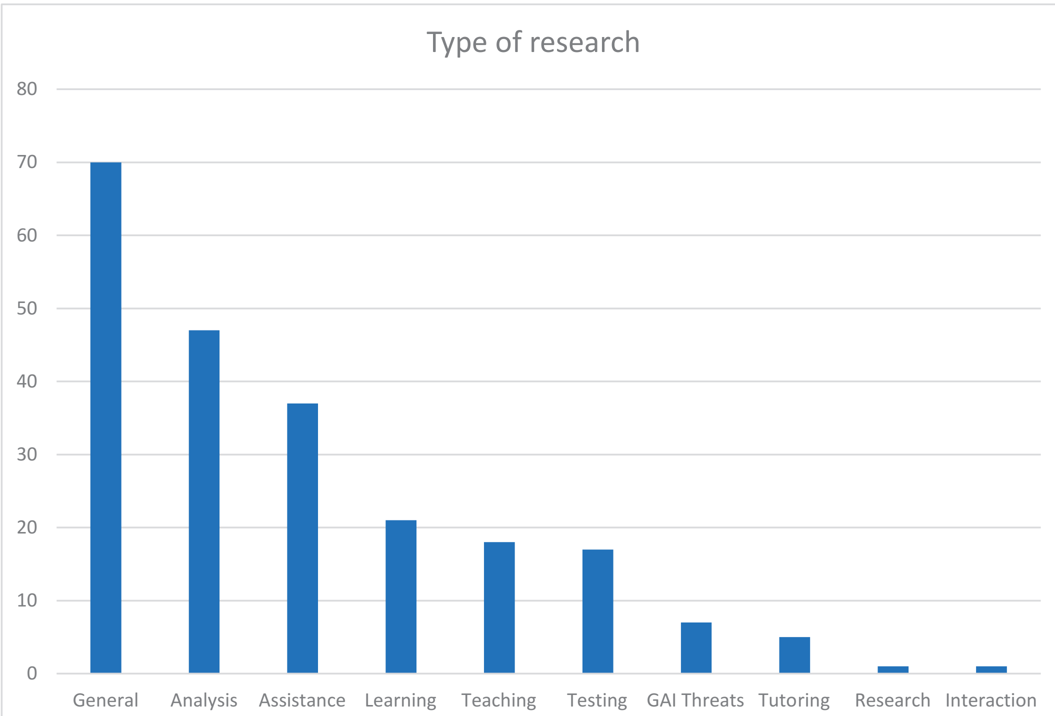


Figure 15. Occurrences of research types in education.

The most prevalent type of research is "General", with 70 occurrences. This suggests a broad range of papers that cover a wide array of topics, methodologies, and perspectives within the domain of GAI in education. Following general papers, "Analysis" emerges as the second most common type, appearing in 47 papers. This indicates a significant emphasis on examining and interpreting data, models, or educational contexts through analytical lenses, allowing for a deeper understanding of the implications and applications of GAI. "Assistance" is the third most frequent type, featured in 37 papers. This underscores the relevance of research aimed at developing AI systems that provide support and guidance to learners or educators. It suggests a strong interest in exploring how GAI can serve as an intelligent assistant in educational settings. "Learning", "Teaching", and "Testing" papers follow with 21, 18, and 17 occurrences, respectively. These types highlight the focus on investigating how GAI can enhance and optimize various aspects of the learning and assessment processes within educational environments. Moreover, "GAI Threats", "Tutoring", "Research", and "Interaction" types make appearances in a more limited number of papers. These occurrences highlight specific areas of interest, such as addressing potential threats associated with GAI in education, exploring the use of AI systems in tutoring scenarios, conducting research studies, and examining interaction dynamics.

5. Conclusions and Future Research

AI has emerged as a revolutionary force in the era of rapid technological advancements, profoundly transforming various aspects of human life. In the field of education, the emergence of GAI has sparked a transformative revolution, influencing areas such as medical and engineering education. GAI's versatile applications include assessment, personalized learning support, and intelligent tutoring systems. The integration of GAI in education has prompted discussions on ethics, academic integrity, and its potential to reshape teaching and learning methods. As GAI continues to advance, thorough research on its applications, implications, and challenges becomes essential in shaping the future of education. This review paper comprehensively analyzed the research on GAI in education, synthesized key findings and insights from 207 publications, and identified research gaps and future directions in the field.

The first step of this work was to conduct a comprehensive examination of the utilization of GAI in education through an in-depth study and content analysis of the literature in the field.

The content analysis has provided significant findings that highlight the transformative impact of GAI in education. Across various domains and contexts, the integration of GAI, including AI chatbots and virtual instructors, offers substantial opportunities for enhancing educational practices and improving learning outcomes. In computer science education, GAI has proven to be a powerful tool in teaching and learning programming concepts. The research papers revealed the potential benefits of AI technologies for creative programming education and the use of AI-powered code generation tools, like Copilot. Engineering education also benefits greatly from GAI's integration, empowering students to generate innovative solutions through advanced chatbots and text-generation models. Cloud-based frameworks and social robots enhance interactive learning experiences, creating a technologically advanced and effective engineering education landscape. Medical education sees promising applications of GAI in fields like medicine, dentistry, pharmacy, and public health education. Nursing education benefits from ChatGPT in enhancing patient care and fostering critical thinking skills among nursing students. In communication education, GAI models showcase potential in journalism, media education, and healthcare communication. The studies demonstrate GAI's effectiveness in enhancing translation accuracy, emphasizing the importance of educating students about responsible AI use and ethical considerations. In academia, GAI offers potential benefits in various academic fields, such as using AI technologies, natural language processing techniques, and LLMs for research. Furthermore, the subsection on "General Education" provided a comprehensive overview of GAI's wider applications in the educational landscape. Diverse research papers explored technology adoption, classroom design, assessment practices, language instruction, ethics, subject-specific education, learning outcomes, student engagement, challenges in AI integration, and innovative pedagogical approaches.

After conducting the content analysis, the next step involved a bibliometric analysis of GAI within the context of education. It included an examination of a total of 437 pre-existing articles and publications retrieved from the Scopus database. After conducting a literature screening and removing duplicates, the analysis comprised a list of 217 publications that focused on GAI in education.

The comprehensive bibliometric analysis of GAI in education yielded several key findings. The research on GAI experienced significant and exponential growth from 2018 to 2023, with a notable surge in papers published in 2023, which was likely due to the popularity and innovation sparked by tools like ChatGPT. In addition, ChatGPT emerged as the dominant AI tool in GAI research within educational contexts, indicating its widespread adoption for intelligent interactions and contextually relevant responses. The interdisciplinary nature of GAI research in education, spanning computer science, engineering education, and medical education, highlights the need for collaboration in addressing complex challenges and potential applications.

While highlighting the promising contributions of GAI integration in education, this review also delved into potential concerns associated with its implementation. This paper navigated the landscape of ethical implications, responsible GAI use, data privacy safeguards, biases, and academic integrity when exploring the assimilation of GAI within educational settings. These multifaceted concerns were considered in various sections of this review paper.

Based on the findings from this review paper, several future research directions can be suggested to advance the field of GAI in educational settings:

- Exploring Transparency: Future research should focus on enhancing the transparency of GAI models. Understanding how AI-generated outputs are generated and providing transparent explanations to users can increase user trust and acceptance of GAI tools in educational settings;
- Addressing Biases and Fairness: As GAI models learn from existing data, they may inherit biases present in the data, potentially leading to biased outputs. Future research should focus on addressing and mitigating biases in GAI tools, especially in educational contexts, to not perpetuate stereotypes or discriminate against certain groups of learners;
- Designing AI-Enhanced Curriculum: Researchers can explore the potential of GAI in generating adaptive and dynamic learning materials that cater to individual learners' needs. AI-powered curriculum design can foster personalized and engaging learning experiences, promoting lifelong learning and skill development;
- GAI in Teacher Professional Development: Research can investigate how GAI tools can assist educators in improving their teaching practices, developing tailored instructional materials, and providing real-time feedback on their performance;
- Collaborative AI in Education: Research the potential of collaborative AI systems in education, where humans and AI work together to achieve educational goals;
- Longitudinal Studies: Conduct longitudinal studies to track the long-term effects of GAI integration in education. These studies can provide insights into the sustained impact of GAI on learning outcomes, retention rates, and academic performance over extended periods;
- Privacy and Data Security: Research should focus on developing robust data protection measures and ensuring that student data is handled responsibly and securely;
- Long-term Impact on Learning Outcomes: Investigate the long-term impact of GAI integration on learning outcomes, academic achievement, and students' problem-solving skills;
- Ethical Considerations and Responsible AI: Further explore the ethical implications of using GAI in education, addressing concerns related to plagiarism, academic integrity, and the potential impact on students' critical thinking skills. Develop guidelines and policies to ensure the responsible and ethical use of GAI technologies in educational settings;
- Student Acceptance and Adoption: Conduct research on students' acceptance and adoption of GAI technology in the learning process. Understand factors influencing their attitudes toward AI-powered tools and identify strategies to enhance student engagement and acceptance;
- Interdisciplinary Collaborations: Foster interdisciplinary collaborations between educators, AI researchers, and policymakers to develop comprehensive frameworks for integrating GAI into education;
- Inclusivity and Accessibility: Explore ways to make GAI-powered educational tools more accessible to diverse learners, including those with disabilities or language barriers;
- AI in Assessment and Evaluation: Investigate the effectiveness of GAI in assessment practices, including essay grading, problem-solving evaluation, and personalized feedback;
- AI and Pedagogy: Investigate the implications of GAI for pedagogical practices, examining how it can complement and enhance traditional teaching methods.

This literature review revealed the immense potential of GAI in reshaping education across various disciplines. While GAI presents exciting opportunities for improving teaching, learning, and educational processes, responsible and ethical usage, as well as addressing potential biases and academic integrity, are crucial aspects that demand attention. The continued exploration of GAI's potential, along with developing comprehensive guidelines and fostering critical thinking skills, will pave the way for a technologically advanced, inclusive, and effective educational landscape.

Author Contributions: Conceptualization, Z.B., C.A. and V.A.; methodology, Z.B., C.A. and V.A. and A.Z.; software, A.Z.; formal analysis, Z.B., C.A. and V.A.; investigation, Z.B., C.A. and V.A.; resources, Z.B. and C.A.; data curation, Z.B., C.A., V.A. and A.Z.; writing—original draft preparation, Z.B., C.A., V.A. and A.Z.; writing—review and editing, Z.B., C.A., V.A. and A.Z.; visualization, Z.B., C.A., V.A. and A.Z.; supervision, Z.B.; project administration, Z.B. All authors have read and agreed to the published version of the manuscript.

Funding: The authors acknowledge the support of the American University of Sharjah under the Open Access Program. This paper represents the opinions of the authors and does not mean to represent the position or opinions of the American University of Sharjah.

Data Availability Statement: Not applicable.

Acknowledgments: The authors express their gratitude to Melvern Moras, an undergraduate student, for his assistance in conducting the bibliometric analysis.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. OpenAI ChatGPT. Available online: https://chat.openai.com (accessed on 28 July 2023).
2. Sullivan, M.; Kelly, A.; McLaughlan, P. ChatGPT in higher education: Considerations for academic integrity and student learning. *J. Appl. Learn. Teach.* **2023**, *6*, 31–40. [CrossRef]
3. Bahrini, A.; Khamoshifar, M.; Abbasimehr, H.; Riggs, R.J.; Esmaeili, M.; Majdabadkohne, R.M.; Pasehvar, M. ChatGPT: Applications, opportunities, and threats. In Proceedings of the 2023 Systems and Information Engineering Design Symposium (SIEDS), Charlottesville, VA, USA, 27–28 April 2023. [CrossRef]
4. Google Bard. Available online: https://bard.google.com/ (accessed on 28 July 2023).
5. Github Copilot. Available online: https://copilot.github.com/ (accessed on 28 July 2023).
6. Hadi, M.U.; Qureshi, R.; Shah, A.; Irfan, M.; Zafar, A.; Shaikh, M.B.; Akhtar, N.; Wu, J.; Mirjalili, S. A Survey on Large Language Models: Applications, Challenges, Limitations, and Practical Usage. *TechRxiv* **2023**. [CrossRef]
7. Baidoo-Anu, D.; Owusu Ansah, L. Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning. Available online: https://ssrn.com/abstract=4337484 (accessed on 28 July 2023).
8. Crompton, H.; Burke, D. Artificial intelligence in higher education: The state of the field. *Int. J. Educ. Technol. High. Educ.* **2023**, *20*, 22. [CrossRef]
9. Sallam, M.; Salim, N.; Barakat, M.; Al-Tammemi, A. ChatGPT applications in medical, dental, pharmacy, and public health education: A descriptive study highlighting the advantages and limitations. *Narra J.* **2023**, *3*, e103. [CrossRef]
10. Hsu, M.; Chan, T.; Yu, C. Termbot: A Chatbot-Based Crossword Game for Gamified Medical Terminology Learning. *Int. J. Environ. Res. Public Health* **2023**, *20*, 4185. [CrossRef]
11. MacNeil, S.; Tran, A.; Mogil, D.; Bernstein, S.; Ross, E.; Huang, Z. Generating diverse code explanations using the gpt-3 large language model. In Proceedings of the 2022 ACM Conference on International Computing Education Research, Lugano, Switzerland, 7–11 August 2022; Volume 2. [CrossRef]
12. Rahman, M.M.; Watanobe, Y. ChatGPT for education and research: Opportunities, threats, and strategies. *Appl. Sci.* **2023**, *13*, 5783. [CrossRef]
13. Kooli, C. Chatbots in education and research: A critical examination of ethical implications and solutions. *Sustainability* **2023**, *15*, 5614. [CrossRef]
14. Markel, J.M.; Opferman, S.G.; Landay, J.A.; Piech, C. GPTeach: Interactive TA Training with GPT Based Students. In Proceedings of the Tenth ACM Conference on Learning @ Scale (L@S '23), Association for Computing Machinery, New York, NY, USA, 20–22 July 2023.
15. Prisma Framework. Available online: http://www.prisma-statement.org/ (accessed on 6 August 2023).
16. Jonsson, M.; Tholander, J. Cracking the code: Co-coding with AI in creative programming education. In Proceedings of the 14th Conference on Creativity and Cognition, Venice, Italy, 20–23 June 2022. [CrossRef]

17. Finnie-Ansley, J.; Denny, P.; Becker, B.A.; Luxton-Reilly, A.; Prather, J. The robots are coming: Exploring the implications of openai codex on introductory programming. In Proceedings of the 24th Australasian Computing Education Conference, Virtual Event, 14–18 February 2022. [CrossRef]
18. Denny, P.; Kumar, V.; Giacaman, N. Conversing with copilot: Exploring prompt engineering for solving cs1 problems using natural language. In Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1, Toronto, ON, Canada, 15–18 March 2023. [CrossRef]
19. Kazemitabaar, M.; Chow, J.; Ma, C.K.T.; Ericson, B.J.; Weintrop, D.; Grossman, T. Studying the effect of AI code generators on supporting novice learners in introductory programming. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, Hamburg, Germany, 23–28 April 2023. [CrossRef]
20. Savelka, J.; Agarwal, A.; Bogart, C.; Sakr, M. Large language models (gpt) struggle to answer multiple-choice questions about code. *arXiv* **2023**, arXiv:2303.08033.
21. Wermelinger, M. Using GitHub copilot to solve simple programming problems. In Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1, Toronto, ON, Canada, 15–18 March 2023. [CrossRef]
22. Becker, B.A.; Denny, P.; Finnie-Ansley, J.; Luxton-Reilly, A.; Prather, J.; Santos, E.A. Programming is hard-or at least it used to be: Educational opportunities and challenges of AI code generation. In Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1, Toronto, ON, Canada, 15–18 March 2023. [CrossRef]
23. Jacques, L. Teaching CS-101 at the Dawn of ChatGPT. *ACM Inroads* **2023**, *14*, 40–46. [CrossRef]
24. Raji, I.D.; Scheuerman, M.K.; Amironesei, R. You can't sit with us: Exclusionary pedagogy in AI ethics education. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, Virtual Event, 3–10 March 2021. [CrossRef]
25. Yilmaz, R.; Yilmaz, F.G.K. The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. *Comput. Educ. Artif. Intell.* **2023**, *4*, 100147. [CrossRef]
26. Cortiñas-Lorenzo, K.; Lacey, G. Toward Explainable Affective Computing: A Review. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**. *accepted*. [CrossRef] [PubMed]
27. Marquez, R.; Barrios, N.; Vera, R.E.; Mendez, M.E.; Tolosa, L.; Zambrano, F.; Li, Y. A perspective on the synergistic potential of artificial intelligence and product-based learning strategies in biobased materials education. *Educ. Chem. Eng.* **2023**, *44*, 164–180. [CrossRef]
28. Humphry, T.; Fuller, A.L. Potential ChatGPT Use in Undergraduate Chemistry Laboratories. *J. Chem. Educ.* **2023**, *100*, 1434–1436. [CrossRef]
29. Sánchez-Ruiz, L.M.; Moll-López, S.; Nuñez-Pérez, A.; Moraño-Fernández, J.A.; Vega-Fleitas, E. ChatGPT Challenges Blended Learning Methodologies in Engineering Education: A Case Study in Mathematics. *Appl. Sci.* **2023**, *13*, 6039. [CrossRef]
30. Nikolic, S.; Daniel, S.; Haque, R.; Belkina, M.; Hassan, G.M.; Grundy, S.; Lyden, S.; Neal, P.; Sandison, C. 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.* **2023**, *48*, 559–614. [CrossRef]
31. Gmeiner, F.; Yang, H.; Yao, L.; Holstein, K.; Martelaro, N. Exploring challenges and opportunities to support designers in learning to co-create with AI-based manufacturing design tools. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, Hamburg, Germany, 23–28 April 2023. [CrossRef]
32. Chen, C.; Olajoyegbe, T.O.; Morkos, B. The imminent educational paradigm shift: How artificial intelligence will reframe how we educate the next generation of engineering designers. In Proceedings of the 2020 ASEE Virtual Annual Conference Content Access, Virtual Event, 22–26 June 2020. [CrossRef]
33. Galdon, F.; Hall, A.; Ferrarello, L. Enhacing abductive reasoning in design and engineering education via probabilistic knowledge; a case study in ai. In Proceedings of the DS 110: 23rd International Conference on Engineering and Product Design Education (E&PDE 2021), VIA Design, VIA University, Herning, Denmark, 9–10 September 2021. [CrossRef]
34. Elfaki, A.O.; Abduljabbar, M.; Ali, L.; Alnajjar, F.; Mehiar, D.; Marei, A.M.; Alhmiedat, T.; Al-Jumaily, A. Revolutionizing Social Robotics: A Cloud-Based Framework for Enhancing the Intelligence and Autonomy of Social Robots. *Robotics* **2023**, *12*, 48. [CrossRef]
35. Konecki, M.; Konecki, M.; Biškupić, I. Using artificial intelligence in higher education. In Proceedings of the 15th International Conference on Computer Supported Education, Prague, Czech Republic, 21–23 April 2023. [CrossRef]
36. Eager, B.; Brunton, R. Prompting higher education towards AI-augmented teaching and learning practice. *J. Univ. Teach. Learn. Pract.* **2023**, *20*, 02. [CrossRef]
37. Chaudhry, I.S.; Sarwary, S.A.M.; El Refae, G.A.; Chabchoub, H. Time to Revisit Existing Student's Performance Evaluation Approach in Higher Education Sector in a New Era of ChatGPT—A Case Study. *Cogent Educ.* **2023**, *10*, 2210461. [CrossRef]
38. Iskender, A. Holy or unholy? Interview with open AI's ChatGPT. *Eur. J. Tour. Res.* **2023**, *34*, 3414. [CrossRef]
39. Strzelecki, A. To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. *Interact. Learn. Environ.* **2023**, *34*, 3414. [CrossRef]
40. Stojanov, A. Learning with ChatGPT 3.5 as a more knowledgeable other: An autoethnographic study. *Int. J. Educ. Technol. High. Educ.* **2023**, *20*, 35. [CrossRef]
41. Pinto, A.S.; Abreu, A.; Costa, E.; Paiva, J. How Machine Learning (ML) is Transforming Higher Education: A Systematic Literature Review. *J. Inf. Syst. Eng. Manag.* **2023**, *8*, 21168. [CrossRef]

42. Waisberg, E.; Ong, J.; Masalkhi, M.; Kamran, S.A.; Zaman, N.; Sarker, P.; Lee, A.G.; Tavakkoli, A. GPT-4: A new era of artificial intelligence in medicine. *Ir. J. Med. Sci.* **2023**, 1–4. [CrossRef]
43. Sevgi, U.T.; Erol, G.; Doğruel, Y.; Sönmez, O.F.; Tubbs, R.S.; Güngor, A. The role of an open artificial intelligence platform in modern neurosurgical education: A preliminary study. *Neurosurg. Rev.* **2023**, *46*, 86. [CrossRef]
44. Šlapeta, J. Are ChatGPT and other pretrained language models good parasitologists? *Trends Parasitol.* **2023**, *39*, 314–316. [CrossRef] [PubMed]
45. Huh, S. Are ChatGPT's knowledge and interpretation ability comparable to those of medical students in Korea for taking a parasitology examination?: A descriptive study. *J. Educ. Eval. Health Prof.* **2023**, *20*, 1–6. [CrossRef]
46. Dolezal, J.M.; Wolk, R.; Hieromnimon, H.M.; Howard, F.M.; Srisuwananukorn, A.; Karpeyev, D.; Ramesh, S.; Kochanny, S.; Kwon, J.W.; Agni, M.; et al. Deep learning generates synthetic cancer histology for explainability and education. *NPJ Precis. Oncol.* **2023**, *7*, 49. [CrossRef] [PubMed]
47. Skalidis, I.; Cagnina, A.; Luangphiphat, W.; Mahendiran, T.; Muller, O.; Abbe, E.; Fournier, S. ChatGPT takes on the European Exam in Core Cardiology: An artificial intelligence success story? *Eur. Heart J.-Digit. Health* **2023**, *4*, 279–281. [CrossRef] [PubMed]
48. Corsello, A.; Santangelo, A. May Artificial Intelligence Influence Future Pediatric Research?—The Case of ChatGPT. *Children* **2023**, *10*, 757. [CrossRef]
49. Lyu, Q.; Tan, J.; Zapadka, M.E.; Ponnatapuram, J.; Niu, C.; Wang, G.; Whitlow, C.T. Translating radiology reports into plain language using chatgpt and gpt-4 with prompt learning: Promising results, limitations, and potential. *arXiv* **2023**, arXiv:2303.09038. [CrossRef]
50. Abdel-Messih, M.S.; Kamel Boulos, M.N. ChatGPT in clinical toxicology. *JMIR Med. Educ.* **2023**, *9*, e46876. [CrossRef]
51. Oh, N.; Choi, G.; Lee, W.Y. ChatGPT goes to the operating room: Evaluating GPT-4 performance and its potential in surgical education and training in the era of large language models. *Ann. Surg. Treat. Res.* **2023**, *104*, 269. [CrossRef]
52. Lee, P.; Bubeck, S.; Petro, J. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. *N. Engl. J. Med.* **2023**, *388*, 1233–1239. [CrossRef]
53. Yang, J.; Por, L.Y.; Leong, M.C.; Ku, C.S. The Potential of ChatGPT in Assisting Children with Down Syndrome. *Ann. Biomed. Eng.* **2023**, 1–3. [CrossRef]
54. Wang, H.; Jia, S.; Li, Z.; Duan, Y.; Tao, G.; Zhao, Z. A Comprehensive Review of Artificial Intelligence in Prevention and Treatment of COVID-19 Pandemic. *Front. Genet.* **2022**, *13*, 845305. [CrossRef] [PubMed]
55. Cascella, M.; Montomoli, J.; Bellini, V.; Bignami, E. Evaluating the feasibility of ChatGPT in healthcare: An analysis of multiple clinical and research scenarios. *J. Med. Syst.* **2023**, *47*, 33. [CrossRef] [PubMed]
56. Gupta, R.; Herzog, I.; Weisberger, J.; Chao, J.; Chaiyasate, K.; Lee, E.S. Utilization of ChatGPT for plastic surgery research: Friend or foe? *J. Plast. Reconstr. Aesthetic Surg.* **2023**, *80*, 145–147. [CrossRef] [PubMed]
57. Masters, K. Ethical use of artificial intelligence in health professions education: AMEE Guide No. 158. *Med. Teach.* **2023**, *45*, 574–584. [CrossRef] [PubMed]
58. Choi, E.P.H.; Lee, J.J.; Ho, M.; Kwok, J.Y.Y.; Lok, K.Y.W. Chatting or cheating? The impacts of ChatGPT and other artificial intelligence language models on nurse education. *Nurse Educ. Today* **2023**, *125*, 105796. [CrossRef]
59. Irwin, P.; Jones, D.; Fealy, S. What is ChatGPT and what do we do with it? Implications of the age of AI for nursing and midwifery practice and education: An editorial. *Nurse Educ. Today* **2023**, *127*, 105835. [CrossRef]
60. Abdulai, A.; Hung, L. Will ChatGPT undermine ethical values in nursing education, research, and practice. *Nurs. Inq.* **2023**, *30*, e12556. [CrossRef]
61. Seney, V.; Desroches, M.L.; Schuler, M.S. Using ChatGPT to teach enhanced clinical judgment in nursing education. *Nurse Educ.* **2023**, *48*, 124. [CrossRef]
62. Ahmed, S.K. The Impact of ChatGPT on the Nursing Profession: Revolutionizing Patient Care and Education. *Ann. Biomed. Eng.* **2023**, 1–2. [CrossRef]
63. Lo, L.S. The CLEAR path: A framework for enhancing information literacy through prompt engineering. *J. Acad. Librariansh.* **2023**, *49*, 102720. [CrossRef]
64. Wang, H. Short sequence Chinese-English machine translation based on generative adversarial networks of emotion. *Comput. Intell. Neurosci.* **2022**, *2022*, 3385477. [CrossRef] [PubMed]
65. Santandreu-Calonge, D.; Medina-Aguerrebere, P.; Hultberg, P.; Shah, M. Can ChatGPT improve communication in hospitals? *Prof. Inf.* **2023**, *32*, e320219. [CrossRef]
66. Pavlik, J.V. Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *J. Mass Commun. Educ.* **2023**, *78*, 84–93. [CrossRef]
67. Feng, Y.; Poralla, P.; Dash, S.; Li, K.; Desai, V.; Qiu, M. The impact of chatgpt on streaming media: A crowdsourced and data-driven analysis using twitter and reddit. In Proceedings of the 2023 IEEE 9th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS), New York, NY, USA, 6–8 May 2023. [CrossRef]
68. Alqahtani, T.; Badreldin, H.A.; Alrashed, M.; Alshaya, A.I.; Alghamdi, S.S.; bin Saleh, K.; Alowais, S.A.; Alshaya, O.A.; Rahman, I.; Al Yami, M.S.; et al. The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research. *Res. Soc. Adm. Pharm.* **2023**, *19*, 1236–1242. [CrossRef] [PubMed]

69. Cotton, D.R.; Cotton, P.A.; Shipway, J.R. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innov. Educ. Teach. Int.* **2023**, 1–12. [CrossRef]
70. Currie, G.M. Academic integrity and artificial intelligence: Is ChatGPT hype, hero or heresy? *Semin. Nucl. Med.* **2023**, *53*, 719–730. [CrossRef]
71. Dergaa, I.; Chamari, K.; Zmijewski, P.; Saad, H.B. From human writing to artificial intelligence generated text: Examining the prospects and potential threats of ChatGPT in academic writing. *Biol. Sport* **2023**, *40*, 615–622. [CrossRef]
72. Perkins, M. Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond. *J. Univ. Teach. Learn. Pract.* **2023**, *20*, 07. [CrossRef]
73. Inamdar, S. Impact of artificial intelligence text generators (AITGs) on libraries. *Libr. Hi Tech News* **2023**. *ahead-of-print*. [CrossRef]
74. Frederick, D.E. ChatGPT: A viral data-driven disruption in the information environment. *Libr. Hi Tech News* **2023**, *40*, 4–10. [CrossRef]
75. Al Ghatrifi, M.O.M.; Al Amairi, J.S.S.; Thottoli, M.M. Surfing the technology wave: An international perspective on enhancing teaching and learning in accounting. *Comput. Educ. Artif. Intell.* **2023**, *4*, 100144. [CrossRef]
76. Geerling, W.; Mateer, G.D.; Wooten, J.; Damodaran, N. ChatGPT has aced the test of understanding in college economics: Now what? *Am. Econ.* **2023**, 05694345231169654. [CrossRef]
77. Skavronskaya, L.; Hadinejad, A.; Cotterell, D. Reversing the threat of artificial intelligence to opportunity: A discussion of ChatGPT in tourism education. *J. Teach. Travel Tour.* **2023**, *23*, 253–258. [CrossRef]
78. Elnozahy, W.A.; El Khayat, G.A. Multi-lang question answering framework for decision support in educational institutes. In Proceedings of the 15th International Conference on Computer Supported Education, Prague, Czech Republic, 21–23 April 2023. [CrossRef]
79. Aydın, N.; Erdem, O.A. A research on the new generation artificial intelligence technology generative pretraining transformer 3. In Proceedings of the 2022 3rd International Informatics and Software Engineering Conference (IISEC), Ankara, Turkey, 15–16 December 2022. [CrossRef]
80. Cooper, G. Examining science education in chatgpt: An exploratory study of generative artificial intelligence. *J. Sci. Educ. Technol.* **2023**, *32*, 444–452. [CrossRef]
81. Su, J.; Yang, W. Unlocking the power of ChatGPT: A framework for applying generative AI in education. *ECNU Rev. Educ.* **2023**, 20965311231168423. [CrossRef]
82. Murugesan, S.; Cherukuri, A.K. The Rise of Generative Artificial Intelligence and Its Impact on Education: The Promises and Perils. *Computer* **2023**, *56*, 116–121. [CrossRef]
83. Hsu, Y.; Ching, Y. Generative Artificial Intelligence in Education, Part One: The Dynamic Frontier. *TechTrends* **2023**, *67*, 603–607. [CrossRef]
84. Mørch, A.I.; Andersen, R. Human-Centred AI in Education in the Age of Generative AI Tools. Available online: https://ceur-ws.org/Vol-3408/short-s2-08.pdf (accessed on 14 July 2023).
85. Gentile, M.; Città, G.; Perna, S.; Allegra, M. Do we still need teachers? navigating the paradigm shift of the teacher's role in the AI era. *Front. Educ.* **2023**, *8*, 1161777. [CrossRef]
86. Jangjarat, K.; Kraiwanit, T.; Limna, P.; Sonsuphap, R. Public Perceptions Towards ChatGPT as the Robo-Assistant. *Online J. Commun. Media Technol.* **2023**, *13*, e202338. [CrossRef]
87. Tlili, A.; Shehata, B.; Adarkwah, M.A.; Bozkurt, A.; Hickey, D.T.; Huang, R.; Agyemang, B. What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learn. Environ.* **2023**, *10*, 15. [CrossRef]
88. Lodge, J.M.; Thompson, K.; Corrin, L. Mapping out a research agenda for generative artificial intelligence in tertiary education. *Australas. J. Educ. Technol.* **2023**, *39*, 1–8. [CrossRef]
89. Chen, X.; Xie, H.; Zou, D.; Hwang, G. Application and theory gaps during the rise of artificial intelligence in education. *Comput. Educ. Artif. Intell.* **2020**, *1*, 100002. [CrossRef]
90. Luo, W.; He, H.; Liu, J.; Berson, I.R.; Berson, M.J.; Zhou, Y.; Li, H. Aladdin's Genie or Pandora's Box for Early Childhood Education? Experts Chat on the Roles, Challenges, and Developments of ChatGPT. *Early Educ. Dev.* **2023**, 2214181. [CrossRef]
91. Kasneci, E.; Seßler, K.; Küchemann, S.; Bannert, M.; Dementieva, D.; Fischer, F.; Gasser, U.; Groh, G.; Günnemann, S.; Hüllermeier, E.; et al. ChatGPT for good? On opportunities and challenges of large language models for education. *Learn. Individ. Differ.* **2023**, *103*, 102274. [CrossRef]
92. Golovianko, M.; Gryshko, S.; Terziyan, V.; Tuunanen, T. Responsible cognitive digital clones as decision-makers: A design science research study. *Eur. J. Inf. Syst.* **2022**, 1–23. [CrossRef]
93. Ye, F.; Bors, A.G. Lifelong teacher-student network learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 6280–6296. [CrossRef]
94. Karadag, I.; Güzelci, O.Z.; Alaçam, S. EDU-AI: A twofold machine learning model to support classroom layout generation. *Constr. Innov.* **2022**, *23*, 898–914. [CrossRef]
95. Vartiainen, H.; Tedre, M. Using artificial intelligence in craft education: Crafting with text-to-image generative models. *Digit. Creat.* **2023**, *34*, 1–21. [CrossRef]
96. Zhai, C.; Wibowo, S.; Cowling, M. An innovative tool for education: An adversarial dialogue system embedded with humor, empathy and culture. In Proceedings of the 2023 11th International Conference on Information and Education Technology (ICIET), Fujisawa, Japan, 18–20 March 2023. [CrossRef]

97. Sarr, J.M.A.; Yannakakis, G.N.; Liapis, A.; Bah, A.; Cambier, C. Djehuty: A mixed-initiative handwriting game for preschoolers. In Proceedings of the 15th International Conference on the Foundations of Digital Games, Bugibba, Malta, 15–18 September 2020. [CrossRef]
98. Bauer, E.; Greisel, M.; Kuznetsov, I.; Berndt, M.; Kollar, I.; Dresel, M.; Fischer, M.R.; Fischer, F. Using natural language processing to support peer-feedback in the age of artificial intelligence: A cross-disciplinary framework and a research agenda. *Br. J. Educ. Technol.* **2023**, *54*, 1222–1245. [CrossRef]
99. Farrokhnia, M.; Banihashem, S.K.; Noroozi, O.; Wals, A. A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innov. Educ. Teach. Int.* **2023**, 1–15. [CrossRef]
100. Howell, B.E.; Potgieter, P.H. What do telecommunications policy academics have to fear from GPT-3? *Telecommun. Policy* **2023**, 102576. [CrossRef]
101. Huang, Y.; Khan, S.M. Advances in AI and machine learning for education research. In *Computational Psychometrics: New Methodologies for a New Generation of Digital Learning and Assessment: With Examples in R and Python*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 195–208. [CrossRef]
102. Ibrahim, H.; Asim, R.; Zaffar, F.; Rahwan, T.; Zaki, Y. Rethinking Homework in the Age of Artificial Intelligence. *IEEE Intell. Syst.* **2023**, *38*, 24–27. [CrossRef]
103. Sung, C.; Dhamecha, T.I.; Mukhi, N. Improving short answer grading using transformer-based pre-training. In Proceedings of the Artificial Intelligence in Education: 20th International Conference, AIED 2019, Chicago, IL, USA, 25–29 June 2019. [CrossRef]
104. Lo, C.K. What is the impact of ChatGPT on education? A rapid review of the literature. *Educ. Sci.* **2023**, *13*, 410. [CrossRef]
105. Ajevski, M.; Barker, K.; Gilbert, A.; Hardie, L.; Ryan, F. ChatGPT and the future of legal education and practice. *Law Teach.* **2023**, 1–13. [CrossRef]
106. Johinke, R.; Cummings, R.; Di Lauro, F. Reclaiming the technology of higher education for teaching digital writing in a post— Pandemic world. *J. Univ. Teach. Learn. Pract.* **2023**, *20*, 01. [CrossRef]
107. Kovačević, D. Use of chatgpt in ESP teaching process. In Proceedings of the 2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH), East Sarajevo, Bosnia and Herzegovina, 15–17 March 2023. [CrossRef]
108. Zhao, T.; Song, T. Establishing a Fusion Model of Attention Mechanism and Generative Adversarial Network to Estimate Students' Attitudes in English Classes. *Teh. Vjesn.* **2022**, *29*, 1464–1471. [CrossRef]
109. Tang, J.; Deng, Y. The Design Model of English Graded Teaching Assistant Expert System Based on Improved B/S Three-Tier Structure System. *Mob. Inf. Syst.* **2022**, *2022*, 4167760. [CrossRef]
110. O'Leary, D.E. An analysis of three chatbots: BlenderBot, ChatGPT and LaMDA. *Intell. Syst. Account. Financ. Manag.* **2023**, *30*, 41–54. [CrossRef]
111. Yan, D. Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Educ. Inf. Technol.* **2023**, 1–25. [CrossRef]
112. Bearman, M.; Ajjawi, R. Learning to work with the black box: Pedagogy for a world with artificial intelligence. *Br. J. Educ. Technol.* **2023**, *54*, 1160–1173. [CrossRef]
113. Fernandez, P. Through the looking glass: Envisioning new library technologies AI-text generators as explained by ChatGPT. *Libr. Hi Tech News* **2023**, *40*, 11–14. [CrossRef]
114. Ray, P.P. ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet Things Cyber-Phys. Syst.* **2023**, *3*, 121–154. [CrossRef]
115. Singh, H.; Singh, A. ChatGPT: Systematic Review, Applications, and Agenda for Multidisciplinary Research. *J. Chin. Econ. Bus. Stud.* **2023**, *21*, 193–212. [CrossRef]
116. Danry, V.; Leong, J.; Pataranutaporn, P.; Tandon, P.; Liu, Y.; Shilkrot, R.; Punpongsanon, P.; Weissman, T.; Maes, P.; Sra, M. AI-generated characters: Putting deepfakes to good use. In Proceedings of the CHI Conference on Human Factors in Computing Systems Extended Abstracts, New Orleans, LA, USA, 29 April–5 May 2022. [CrossRef]
117. Ali, S.; DiPaola, D.; Lee, I.; Hong, J.; Breazeal, C. Exploring generative models with middle school students. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 8–13 May 2021. [CrossRef]
118. Adams, T.; Jameel, S.M.; Goggins, J. Education for Sustainable Development: Mapping the SDGs to University Curricula. *Sustainability* **2023**, *15*, 8340. [CrossRef]
119. Emenike, M.E.; Emenike, B.U. Was This Title Generated by ChatGPT? Considerations for Artificial Intelligence Text-Generation Software Programs for Chemists and Chemistry Educators. *J. Chem. Educ.* **2023**, *100*, 1413–1418. [CrossRef]
120. Agathokleous, E.; Saitanis, C.J.; Fang, C.; Yu, Z. Use of ChatGPT: What does it mean for biology and environmental science? *Sci. Total Environ.* **2023**, *888*, 164154. [CrossRef]
121. Hallsworth, J.E.; Udaondo, Z.; Pedrós-Alió, C.; Höfer, J.; Benison, K.C.; Lloyd, K.G.; Cordero, R.J.; de Campos, C.B.; Yakimov, M.M.; Amils, R. Scientific novelty beyond the experiment. *Microb. Biotechnol.* **2023**, *16*, 1131–1173. [CrossRef]
122. Megahed, F.M.; Chen, Y.; Ferris, J.A.; Knoth, S.; Jones-Farmer, L.A. How generative AI models such as chatgpt can be (mis) used in spc practice, education, and research? an exploratory study. *Qual. Eng.* **2023**, 1–29. [CrossRef]
123. Ivanov, S.; Soliman, M. Game of algorithms: ChatGPT implications for the future of tourism education and research. *J. Tour. Futures* **2023**, *9*, 214–221. [CrossRef]
124. Fergus, S.; Botha, M.; Ostovar, M. Evaluating academic answers generated using ChatGPT. *J. Chem. Educ.* **2023**, *100*, 1672–1675. [CrossRef]

125. Siche, R.; Siche, N. El modelo de lenguaje basado en inteligencia artificial sensible-ChatGPT: Análisis bibliométrico y posibles usos en la agricultura y pecuaria. *Sci. Agropecu.* **2023**, *14*, 111–116. [CrossRef]
126. Uddin, S.J.; Albert, A.; Ovid, A.; Alsharef, A. Leveraging ChatGPT to Aid Construction Hazard Recognition and Support Safety Education and Training. *Sustainability* **2023**, *15*, 7121. [CrossRef]
127. Wolfert, R.; van Nederveen, S.; Binnekamp, R. Fit for Purpose Building Information Modelling and Systems Integration (BIMSI) for Better Construction Projects Management. *J. Mod. Proj. Manag.* **2022**, *10*, 174–187. [CrossRef]
128. Wu, R.; Yu, Z. Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. *Br. J. Educ. Technol.* **2023**. [CrossRef]
129. Muñoz, S.A.S.; Gayoso, G.G.; Huambo, A.C.; Tapia, R.D.C.; Incaluque, J.L.; Aguila, O.E.P.; Cajamarca, J.C.R.; Acevedo, J.E.R.; Rivera, H.V.H.; Arias-Gonzáles, J.L. Examining the Impacts of ChatGPT on Student Motivation and Engagement. *Soc. Space* **2023**, *23*, 1–27.
130. Crawford, J.; Cowling, M.; Allen, K. Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *J. Univ. Teach. Learn. Pract.* **2023**, *20*, 02. [CrossRef]
131. Shoufan, A. Exploring Students' Perceptions of CHATGPT: Thematic Analysis and Follow-Up Survey. *IEEE Access* **2023**, *11*, 38805–38818. [CrossRef]
132. Gardner, D.E.; Giordano, A.N. The Challenges and Value of Undergraduate Oral Exams in the Physical Chemistry Classroom: A Useful Tool in the Assessment Toolbox. *J. Chem. Educ.* **2023**, *100*, 1705–1709. [CrossRef]
133. Hemachandran, K.; Verma, P.; Pareek, P.; Arora, N.; Rajesh Kumar, K.V.; Ahanger, T.A.; Pise, A.A.; Ratna, R. Artificial Intelligence: A Universal Virtual Tool to Augment Tutoring in Higher Education. *Comput. Intell. Neurosci.* **2022**, *2022*, 1410448. [CrossRef]
134. Limo, F.A.F.; Tiza, D.R.H.; Roque, M.M.; Herrera, E.E.; Murillo, J.P.M.; Huallpa, J.J.; Flores, V.A.A.; Castillo, A.G.R.; Peña, P.F.P.; Carranza, C.P.M.; et al. Personalized tutoring: ChatGPT as a virtual tutor for personalized learning experiences. *Soc. Space* **2023**, *23*, 293–312.
135. Alnaqbi, N.M.; Fouda, W. Exploring the Role of ChatGPT and social media in Enhancing Student Evaluation of Teaching Styles in Higher Education Using Neutrosophic Sets. *Int. J. Neutrosophic Sci.* **2023**, *20*, 181–190. [CrossRef]
136. Jeon, J.; Lee, S. Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Educ. Inf. Technol.* **2023**, 1–20. [CrossRef]
137. Dijkstra, R.; Genç, Z.; Kayal, S.; Kamps, J. Reading Comprehension Quiz Generation Using Generative Pre-Trained Transformers. 2022. Available online: https://intextbooks.science.uu.nl/workshop2022/files/itb22_p1_full5439.pdf (accessed on 16 July 2023).
138. Pataranutaporn, P.; Leong, J.; Danry, V.; Lawson, A.P.; Maes, P.; Sra, M. AI-generated virtual instructors based on liked or admired people can improve motivation and foster positive emotions for learning. In Proceedings of the 2022 IEEE Frontiers in Education Conference (FIE), Uppsala, Sweden, 8–11 October 2022. [CrossRef]
139. Shu, D.; Doss, C.; Mondschein, J.; Kopecky, D.; Fitton-Kane, V.; Bush, L.; Tucker, C. A pilot study investigating STEM learners' ability to decipher AI-generated video. In Proceedings of the 2021 ASEE Virtual Annual Conference, Virtual Event, 13–15 September 2021.
140. Wardat, Y.; Tashtoush, M.A.; AlAli, R.; Jarrah, A.M. ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia J. Math. Sci. Technol. Educ.* **2023**, *19*, em2286. [CrossRef] [PubMed]
141. Halaweh, M. ChatGPT in education: Strategies for responsible implementation. *Contemp. Educ. Technol.* **2023**, *15*, ep421. [CrossRef] [PubMed]
142. Lim, W.M.; Gunasekara, A.; Pallant, J.L.; Pallant, J.I.; Pechenkina, E. Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *Int. J. Manag. Educ.* **2023**, *21*, 100790. [CrossRef]
143. Hwang, G.; Chen, N. Editorial Position Paper. *Educ. Technol. Soc.* **2023**, *26*, xviii. [CrossRef]
144. Yu, H.; Guo, Y. Generative artificial intelligence empowers educational reform: Current status, issues, and prospects. *Front. Educ.* **2023**, *8*, 1183162. [CrossRef]
145. Dwivedi, Y.K.; Kshetri, N.; Hughes, L.; Slade, E.L.; Jeyaraj, A.; Kar, A.K.; Baabdullah, A.M.; Koohang, A.; Raghavan, V.; Ahuja, M.; et al. "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *Int. J. Inf. Manag.* **2023**, *71*, 102642. [CrossRef]
146. Javed, R.T.; Nasir, O.; Borit, M.; Vanhée, L.; Zea, E.; Gupta, S.; Vinuesa, R.; Qadir, J. Get out of the BAG! Silos in AI ethics education: Unsupervised topic modeling analysis of global AI curricula. *J. Artif. Intell. Res.* **2022**, *73*, 933–965. [CrossRef]
147. Karaali, G. Artificial Intelligence, Basic Skills, and Quantitative Literacy. *Numeracy* **2023**, *16*, 9. [CrossRef]
148. Lawrie, G. Establishing a delicate balance in the relationship between artificial intelligence and authentic assessment in student learning. *Chem. Educ. Res. Pract.* **2023**, *24*, 392–393. [CrossRef]
149. Robertson, L.; Muirhead, B. Where do we go from here? An interdisciplinary exploration of leveraging new technologies in education. In Proceedings of the International Conference on Society and Information Technologies (ICSIT0), Orlando, FL, USA, 8–11 March 2022. [CrossRef]
150. Álvarez-Álvarez, C.; Falcon, S. Students' preferences with university teaching practices: Analysis of testimonials with artificial intelligence. *Educ. Technol. Res. Dev.* **2023**, 1–16. [CrossRef]
151. Van Eck, N.J.; Waltman, L. *VOSviewer Manual: Manual for VOSviewer Version 1.6.15*; Centre for Science and Technology Studies (CWTS) of Leiden University: Leiden, The Netherlands, 2020.

152. Januszewski, A.; Żółtowski, D. Emerging ICT for Sustainable Development. Research Concept of Literature Analysis. In Proceedings of the Americas Conference on Information Systems (AMCIS) 2023, Panama City, Panama, 10–12 August 2023.
153. Van Eck, N.J.; Waltman, L. Text mining and visualization using VOSviewer. *arXiv* **2011**, arXiv:1109.2058.
154. OpenAI Codex. Available online: https://openai.com/blog/openai-codex (accessed on 28 July 2023).

sustainability

MDPI

Article

The Role of Dynamic Geometry Software in Teacher–Student Interactions: Stories from Three Chinese Mathematics Teachers

Fangchun Zhu * and Binyan Xu

Asian Centre for Mathematics Education, East China Normal University, Shanghai 200241, China
* Correspondence: fczhu@math.ecnu.edu.cn

Abstract: Teacher–student interactions are critical for effective mathematics teaching and learning. Certain technologies, especially dynamic geometry software (DGS), can have a great impact on interactions between teachers and students. In Chinese mathematics educational environments, we selected three mathematics teachers as examples to analyze how DGS impacts their method of teaching and feedback and what role it plays in their interactions with students. Based on the data, we found mathematics to be the main focus for Chinese teachers, even though DGS was widely used in their lessons. DGS was mainly used as a projector to present geometric properties. This implies that the interactivity of the DGS was simply used for pressing buttons on the screen during lessons. Although teachers may have used DGS for many years, they still lack sufficient skills to integrate this software into mathematics teaching.

Keywords: teacher–student interactions; dynamic geometry software; Chinese mathematics lessons; technology

Citation: Zhu, F.; Xu, B. The Role of Dynamic Geometry Software in Teacher–Student Interactions: Stories from Three Chinese Mathematics Teachers. *Sustainability* **2023**, *15*, 7660. https://doi.org/10.3390/su15097660

Academic Editors: Jian-Hong Ye, Yung-Wei Hao, Yu-Feng Wu and Savvas A. Chatzichristofis

Received: 8 March 2023
Revised: 4 May 2023
Accepted: 5 May 2023
Published: 6 May 2023

1. Introduction

Teacher–student interactions form a critical part of classroom teaching through posing questions and giving feedback, affecting students' learning processes [1]. For a teacher, a student's answers to questions indicate their understanding and knowledge surrounding a certain topic [2]. Additionally, it is also important that teachers ask suitable questions in learning environments with the aid of technology to help students learn mathematics [3]. Technology gives students more opportunities to develop their mathematical skills [4] via their interactions with teachers. For example, the use of technology might encourage students to ask worthwhile questions that may not be posed by teachers [5]. Additionally, answers to these questions can be a catalyst for improving students' problem-solving strategies when their original strategies are unsuitable [6].

Dynamic geometry software (DGS) is the most commonly used technology in mathematics teaching and learning. Its main function, the dragging mode [7], is frequently used by teachers to explain geometry. Studies on this software often focus on the design of different tasks for teachers that use DGS [8,9]. However, few studies focus on what kinds of questions teachers pose to students via the use of dynamic geometry software [3]. Within the situation created by DGS, teachers need to think about designing suitable questions in order to allow students to improve their mathematical knowledge. In this study, we focused on both questions and feedback regarding the interactions between teachers and students using DGS in the classroom.

2. Literature Review

2.1. Dynamic Geometry Software in Mathematics Education

Currently, teachers use different kinds of software to support their teaching process. These kinds of software are designed to teach mathematics as clearly as possible and show the results of mathematical processes [10]. They all have these basic features, such as primitive objects (points, lines, circles, etc.), basic construction tools (parallel, perpendicular, etc.), transformations, and measuring and dragging capabilities [3]. DGS has been used for almost three decades, and many studies have demonstrated its contribution to improving learning and teaching mathematics [11–13] via its dragging mode, which is a problem-solving tool with a number of applications, including construction, searching for commonalities, conjecturing, and proving and refuting conjectures [14,15]. When students drag geometric objects with DGS, the geometrical properties of the figure defined in its construction are preserved.

Consequently, teachers must develop new methods for this new learning environment, in which students can have direct interactions with DGS. Previously, teachers had to directly teach students mathematical concepts [16], whereas, with DGS, teachers can let students independently explore and experiment [17]. Despite this, in most lessons, teachers always reveal answers using DGS without letting students explore solutions independently [18]. Students can only observe the outcomes found by teachers using DGS and have no opportunities to question whether the teachers are actually correct [19].

Because of the complex processes of interaction with this software, a teacher may find that students' experiences of interacting with this software can lead to unexpected outcomes [7].

2.2. Framework

2.2.1. Roles of DGS in Mathematics Teaching and Learning: Amplifier and Reorganizer

In this paper, we aimed to explore the roles that DGS plays during the interactions between teachers and students in geometry tasks. Pea [20] found two distinct roles of DGS in mathematics teaching: amplification and reorganization. According to Pea, technologies are always seen as an amplifier that can help users to solve tasks more efficiently. However, this is not sufficient for mathematical learning and teaching. Therefore, users need to make changes when solving tasks with DGS [21]. Over the years, the use of DGS has begun to change from being a visual amplifier to a reorganizer of the conceptual understanding of learners [22]. Based on these two roles, some researchers have proposed four different roles that DGS plays in mathematics tasks [8,9]:

1. DGS is used to facilitate the material aspects of a task without changing it conceptually. For example, a triangle is constructed using the midpoints of its sides and then its medians.
2. Alternatively, DGS is used as a visual amplifier in the task of identifying properties. It is easier to observe that three midlines of one triangle always intersect at one point via dragging with DGS than by using a pencil and paper.

These two roles can be seen as amplifiers in mathematics tasks. DGS also offers new tasks with new and improved standard geometric knowledge that cannot be achieved without its use.

3. DGS can modify the solving strategies of a task because of its powerful functions, which may make the tasks more difficult, similar to creating a parallelogram without using parallel line tools in DGS.
4. The task itself takes its meaning or its 'raison d'être' from Cabri. These two roles show that DGS can be used to generate new tasks in mathematics.

2.2.2. Initial Categories of Focusing Questions

Teachers, students, and mathematics are the foundations of mathematics classes [3,23,24], but in the 21st century, technology is becoming an increasingly important aspect of our daily lives. The emergence of technology compels us to reconsider the relationship between these elements. Tall [10] proposed a tetrahedron model, placing technology as one of the most important elements in mathematical teaching and learning (Figure 1). This model indicates that teachers may either focus on technology and give students a little opportunity for mathematical learning, or they may focus mostly on mathematics and not provide any technological instructions at all [25,26].

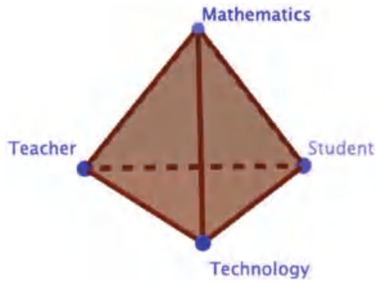
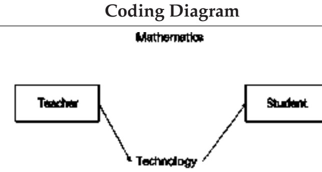
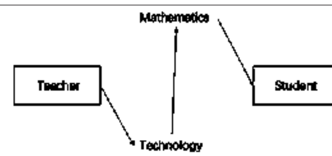
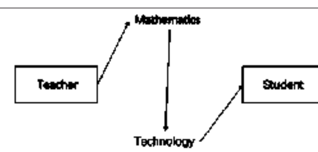
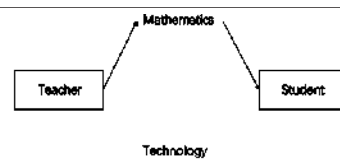


Figure 1. Didactic tetrahedron.

Because of the change in teaching methods caused by DGS, teacher–student interactions can show a different format. We can use Table 1 to describe the interactions between teachers, students, mathematics, and DGS in our lessons.

Table 1. Four initial foci of questions.

Focus	Description	Example Questions	Coding Diagram
Focus on technology.	Teacher's question or feedback is only about the technology.	Try to drag and click the circle button and draw a circle on your screen.	
Focus on technology to notice mathematics.	Teacher's question or feedback is related to a technological action, which is related to a mathematical idea.	If you select the center of the circle you have drawn and move it, what do you see?	
Focus on mathematics with the use of technology.	Teacher's question or feedback is about mathematics, and they assume that students need to use technology to find answers.	Can you make a circle that circumscribes the triangle given on your screen?	
Focus on mathematics.	Teacher's question or feedback are only about mathematics. Additionally, students can answer them without the help of technology.	Try to conclude ways to prove the drawn triangles.	

From this table, four different routes can be identified to show how teachers interact with their students based on mathematics and technology [3]. The route of teacher–mathematics–student means that teachers think that mathematics knowledge is the most important and overlook the technology. In the route of teacher–technology–student, teachers may need to explain the function of the technology they use, and at those times, they do not pay attention to the mathematics. Because of the development of technology, in the routes of teacher–technology–mathematics–students, and teacher–mathematics–technology–students, teachers consider ways to integrate technology to help their students learn mathematics effectively. Sometimes they may want their students to understand how to use technology, so they may pose some mathematics questions as examples to teach. In other cases, teachers may want students to know how to solve mathematical problems with the help of technology.

3. Research Questions

In this study, we addressed the following questions:

1. What technology-specific actions do mathematics teachers focus on when using DGS in mathematics lessons during teacher–student interactions?
2. What roles does DGS play as an amplifier and reorganizer during teacher–student interactions?

4. Methodology

In this section, we describe the criteria for selecting the teachers; their basic information, resources, and tasks; and further explain the data collection and analysis process.

To address RQ1, we paid attention to lesson videos, analyzing the kinds of questions that teachers posed to their students and the feedback given after students answered the questions.

To address RQ2, we focused on how DGS appears as a technological resource to support teachers' interactions with students during lessons.

4.1. Participants

Considering the multiple factors impacting teachers' use of DGS in the classroom, the following criteria were considered: (1) teachers' experience (both experienced and novice teachers); (2) teachers' willingness to use DGS in their lessons; and (3) teachers' experience of using DGS. Based on these criteria and to address our research questions, we decided to choose three Chinese mathematics teachers with different teaching experiences of DGS (two experienced teachers and one novice teacher), who were all willing to use DGS and other technological resources.

Teachers ZH and J were the two most experienced teachers from the same school and the leading mathematics teachers in their respective grades. Tables 2 and 3 present the tasks used by the teachers during this study.

Table 2. The task used by Mr. ZH.

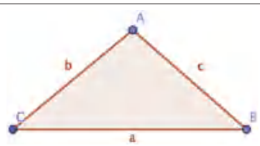
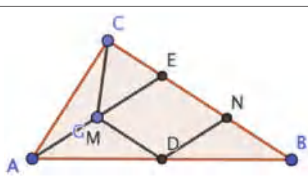
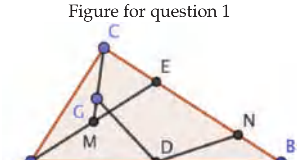
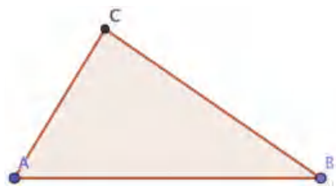
Task	
	Mr. ZH designed one diagram of a triangle whose side lengths AC and BC were determined (AC = 4, BC = 6). Point A could be moved, while points B and C were stable. From the diagram, the trace of point A is similar to a circle. The main question: please find the relationship between the three sides of the triangle. (Here the lowercase letters a, b, and c mean side BC, side AC, and side AB)

Table 3. The task used by Madame J.

Task	
Figure for question 1 Figure for question 2	In the picture, there is a right triangle ABC, ∠ACB = 90°, point D is the midpoint of AB, point E is on segment BC, AE = BE, point M is the midpoint of AE, point G is on segment CM, and point N is on segment BC, which makes ∠GDN = ∠AEB Question 1: as in picture 1, if point G coincides with point M, this proves that quadrilateral DMEN is a diamond. Question 2: as in picture 2, when point G does not coincide with point M and point C, prove GD = DN

Teacher Y was a novice teacher from another school who needed support from colleagues to help her use DGS in the classroom. The task used in her lesson is shown in Table 4. All the lessons observed in this research were held with grade 7 students.

Table 4. The task used by Mrs. Y.

Task	
	Task 1: DGS is used to construct a new triangle, which is congruent with the given triangle ABC, in which AB = 7, AC = 4, BC = 6. In this triangle, point A and B can be dragged, while point C is stable. When point A is dragged, the whole triangle moves without changing its type and size. When point B is dragged, point A does not move. The trajectory of point B is a circle whose center is point A and radius is the length of AB.

4.2. Data Collection and Analysis

According to the research objective, some details from the pictures and interviews may be ignored. The first author observed all three teachers' lessons with DGS and made lesson videos (40 min for each lesson). The following were the main elements that we focused on: the types of tasks (with DGS); the teacher's interactions with students [27]; the organization and configuration of classes; and potential comments made by both teachers and students [28].

Due to the research objective and for ethical reasons, the videos were focused on the teachers, and only the teachers' faces were shown in order to protect their privacy.

These videos were reanalyzed based on the role of DGS and the focus on teachers' questions and feedback. We first cut each lesson video into short teaching episodes according to the different mathematics tasks used in the lessons.

Then, the interactions between the teachers and students were analyzed to determine the focus of the teachers when using DGS in the classroom. (1) If the question or piece of feedback focused on the technology, then it could be classified as the route "teacher–technology–student". (2) If the question or piece of feedback focused on mathematics, then it could be classified as the "teacher–mathematics–student" route. (3) If the question or feedback focused on how students use technology to solve mathematics tasks, it could be classified as the "teacher–mathematics–technology–student" route. (4) If the question or feedback focused on using mathematics contents to present the function of the technology, it could be classified as the "teacher–technology–mathematics–student" route.

To answer RQ2, we analyzed the video recordings according to the different roles that DGS plays in teacher–student interactions (amplifier or generator), and we examined the interviews with these teachers to determine their uses of and opinions on DGS in teaching mathematics. The audio from all the interviews was recorded. First, we transcribed all the audio recordings and cut them into small sections according to the interview questions. Then, all the teachers' answers were analyzed sentence by sentence. We focused on the keywords that applied to teaching methods involving DGS (one example is presented in Appendix A).

5. Results

In this section, we present the interactions involved in the above task to analyze which teachers' questions and feedback points focused on when DGS was used in the classroom and the roles that DGS played in these interactions.

5.1. Stories from Mr. ZH's Class

Based on the task situation (Table 2), Mr. ZH posed the following questions for discussion:

1. What is the maximum length of side c? What is the relationship between sides a, b, and c? What shape is it now?

2. What is the minimum length of side c? What is the relationship between sides a, b, and c? What shape is it now?
3. Based on the conclusions of questions 1 and 2, what you can find?

The students needed to find the mathematical relationships between the three sides of the triangles. Mr. ZH did not expect students to use the software to find the answers; the students just observed what occurred on the screen when the software was running.

Before the discussion, Mr. ZH let the students watch the screen and use DGS as a projector to present the task and diagram. During the discussion, Mr. ZH did not directly interact with his students but acted as a listener. After students finished their work, Mr. ZH continued the lesson with the following question, as presented in Table 5: "OK, the first one, what is the maximum of side C?"

Table 5. Teacher–student interactions in Mr. ZH's lesson.

Questions and Feedback	Focus
T: "OK, the first one, what is the maximum of side c?"	Only focused on mathematics.
St 1: "The maximum is 10" T: "OK, maximum is 10. I have said in the video, so the maximum is 10"	The teacher's feedback just repeated what the student answered. The answer only focused on mathematics.
T: "Now what is the relation?"	Only focused on mathematics.
St 1:" They are all on the same line." T: "Good, he said on the same line."	The teacher's feedback just repeated what the student answered. The answer only focused on mathematics
T: "Let us see, this is position. OK, now c is 10 and they are all on the same line."	The teacher moved the point on the screen at that time. The feedback focused on technology to notice mathematics.
T: "But what is the mathematical relation between the three sides?"	Only focused on mathematics.
St 1: "b + c = a" T: "b + c = a, good, sit down. b + c = a, we can also say a − b = c."	The teacher's feedback just repeated what the student answered. The answer only focused on mathematics.

One student (St1) was asked to answer the question. From Table 5, both Mr. ZH's questions and the student's answers focused more on mathematics. Sometimes, Mr. ZH tried to let the students observe the mathematical content with the help of DGS, but DGS was only used by Mr. ZH (Figure 2) and not his students. This implies that DGS was more like an amplifier that was used to enlarge the geometry properties and support the teachers' explanations. During this time, the students still did not truly explore knowledge with the help of DGS. When the teachers let the students solve the task together or discuss the questions that they posed, the teachers acted as listeners and often did not participate in the students' discussions. Additionally, although they may have used DGS to teach, it did not play an important role during discussions.

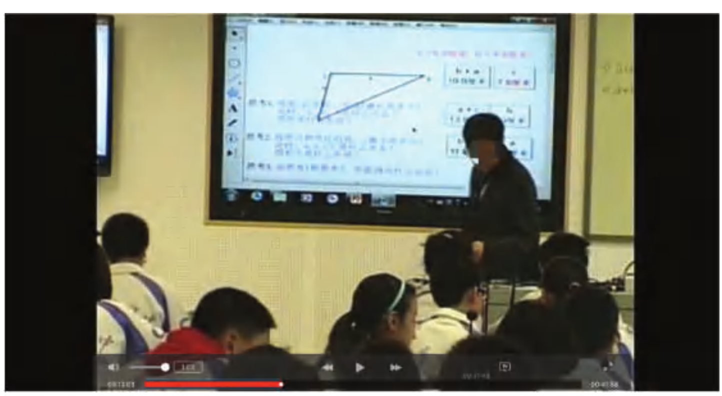


Figure 2. Mr. ZH operated DGS during teacher–student interactions. In this figure, the Chinese terms mean: question 1, please find what is the smallest value of a + b; question 2, please find what is the largest value of a + b; and question 3, what conclusion can you make?

The interaction process was supported by Mr. ZH's preference for using DGS in mathematics education. In the interview, Mr. ZH told us he generally always used DGS to explain complex tasks; for him, it was a personal preference. Because of examinations, Mr. ZH could not allocate the students too much time to use DGS for exploring mathematical concepts. As he stated, in China, students need to pass exams without the help of any software.

5.2. Stories from Madame J's Class

Here, we describe another teacher's classroom interactions through the use of DGS. As mentioned above, Madame J was familiar with DGS in mathematics education. Sometimes, she searched for different tasks on the internet or in some books to learn how to effectively use DGS in the classroom. We chose two examples of teacher–student interactions (Tables 6 and 7) according to the same task (Table 3). One example shows that a student had the incorrect answers and the other one shows that the student had the correct answer.

Table 6. Teacher–student interactions in Madame J's lesson.

Questions and Feedback	Focus
T: "You need to justify what type is the quadrilateral."	Focused on mathematics.
St1: "It is diamond." T: "Diamond. How can you know it?"	The teacher's feedback repeated what this student answered. Additionally, the question let this student justify why they believed that quadrilateral DMEN was a diamond. Therefore, the feedback point and question focused on mathematics.
St1: "First point G and M are coincided." T: "Yes"	The teacher gave positive feedback. They did not focus on mathematics or technology.
St1: "Then, DM, point D is the midpoint of AB, point M is the midpoint of AE, so DM the median line of this triangle." T: "DM is the median line of triangle ABE. You know it is median line, then"	The teacher repeated the conclusion this student made. It still focused on mathematics.
St1:" Then DM is parallel to BE T: DM is parallel to BE"	The teacher repeated the conclusion this student made. It still focused on mathematics.
St1: "So angle MDN is add angle END is 180" T: "It is this angle, angle MDN and angle END, angle MDN plus angle END is 180"	The teacher wrote down some marks through DGS on the screen. The feedback focused on technology to notice mathematics
St1: "Then, it said, angle GEN is equal to angle MDN and AEB" T: "Angle MDN is equal to this angle, angle GEN. It means this angle is equal to angle GEN."	The teacher wrote down some marks through DGS on the screen. The feedback focused on technology to teach mathematics
St1: "So angle END plus angle GEN is 180, then DN is parallel to ME" T: "Why you justify DN is parallel to ME? St1: So NDME is parallelogram"	This question only focused on mathematics.

In her lessons, Madame J always used an iPad as one of her resources, making her lessons a little different than those of other Chinese teachers. As shown in Table 6, Madame J did not pose many questions; instead, she repeated the student's answers or gave feedback such as "ok", and "yes" to let the student continue their explanation. From some critical terms in her questions, such as "diamond" and "parallel", we can see that these questions were more focused on mathematics. Therefore, DGS, similar to the situation in Mr. ZH's lesson, was used to project and display diagrams for the students to observe. They did not need to operate the software to help them find the correct answer (Figure 3).

The next interaction (Table 7) occurred when St1 failed to provide an answer to the question. Madame J asked another question to continue the explanation.

Table 7. Another teacher–student interaction in Madame J's lesson.

Questions and Feedback	Focus
St2: "In the conditions, AE is equal to BE T: AE is equal to BE, what it means?"	The teachers' question focused on mathematics.
St2: "So we know MD is the median line. Then, MD is 1/2 BE" T: "Ok, the median line can help us know not only this one. However, also DM is equal to 1/2 BE."	The teacher explained why this student arrived at their conclusion. The feedback focused on mathematics.
St2: "Because point M and N, M is the midpoint of AE, then, so, it is also useful, it means AM is equal to ME and 1/2AE T: Equal to 1/2AE"	The feedback focused on mathematics
St2: "Then, AE is equal to BE, so MD is equal to ME" T: "MD is equal to " St2: "ME" T: "So now, when AE is equal to BE, this condition is also useful, AE is equal to BE, another one is ME is equal to, we need another condition, it is ME is equal to 1/2 AE. Ok, from these two conditions, we can find ME is equal to DM, then it can be a diamond"	After this student finished their answer, the teacher made a conclusion that focused on mathematics.

Figure 3. Students did not need to operate the software.

Compared with the first interaction, Madame J did not change her teaching strategies, despite the fact that the first students failed to answer her question. She gave both of them enough time to provide their own explanations without any interruptions. However, there were no terms related to DGS mentioned during the interaction. The main focus was only on mathematics. No explanation or method was presented during the interaction. Madame J simply wrote the student's answers on the screen or whiteboard (Figure 4).

Figure 4. Madame J wrote down students' answers on the screen and whiteboard.

DGS was used to project images since Madame J preferred this technology. However, she believed that using DGS requires teachers' guidance and control in lessons.

5.3. Stories from Mrs. Y's Class

Mrs. Y did not have as much teaching experience with DGS compared to the other teachers in this study. Therefore, it was interesting for us to analyze her interactions with students using DGS. Additionally, it helped us to paint a clearer picture of teachers' practices using DGS in mathematics education. The following interaction (Table 8) shows how the students participated in the classroom when Mrs. Y operated DGS and explained what happened.

Table 8. Teacher–student interaction in Mrs. Y's lesson.

Questions and Feedback	Focus
T:"Let's draw another triangle whose sides are also 4, 7, and 6. Then, we can overlap the two triangles and see whether the two triangles coincided, if they coincided, then the two triangles are, congruent. How to draw it?"	The teacher wanted the students to draw a triangle with the help of DGS. Her question and statement focused on mathematics with the use of technology.
St 1: "First draw one base side" T: "First draw one base side. [teacher operated the computer, choose the tool "segment"] We can find the tool "segments with given length" in the tool "lines". Then, here, we draw a segment, 7, right? Here we tap 7. [teacher clicked the tool, and tapped "7"] Now you can see a segment and its length is 7. Ok, then, how to draw this point on side AC, that is side DF?"	The teacher used DGS to show the students the whole drawing process. Her question focused on mathematics with use of technology.
St 2: "Use compass" T: "Use compass, it means the radius of this circle is?"	The teacher's feedback just repeated what the student answered. It focused on mathematics.
All St: "4" T: "4, which one is its center?"	Focused on mathematics.
St 3: "D" T: "D, ok, now we can find circle. [teacher operated the computer, find the tool "circle"] Then, you can click it, it showed "one center and one point". We need to find the circle and radius, it means the radius is determined, click it. [teacher clicked the tool "circle with center and radius", then looked at the students] Ok, now point D is the center, click it. What is the radius?"	The teacher used DGS to show the students the whole drawing process. Her statements focused on mathematics with use of technology, while one question focused on mathematics.
All St: "4" T: "Then, here tap "4", then you can see a circle. [teacher tapped "4", the screen showed one circle whose center is D and radius is 4] Ok, then is this side, its length is 6. Ok, its center is point E, what is the radius? 6. [teacher clicked "circle with center and radius", tapped 6] Ok, now you can find the intersect point of the two circle?"	The teacher used DGS to show the students the whole drawing process. Her statements focused on mathematics with use of technology, while the question focused on mathematics.
St 4: "Two" T: "Yes, two, it's the same with the two points up or down"	Focused on mathematics.
St 5: "Yes, it is same" T: "Ok, same, here I just draw one, this point, which is the intersect point?"	Focused on mathematics.
St 6:" F" T: "F, it is F here, then, we can draw this triangle. [teacher clicked the intersect point] Link these three points, DEF. [teacher operated the mouse, linked point D, E, F] This triangle, ok, these lines are supportive, we can delete them."	The teacher used DGS to show the students the whole drawing process. Her statements focused on mathematics with the use of technology.

According to the task situation (Table 4), the students were asked to construct a triangle with three given sides using DGS. This kind of task is seldom used in Chinese mathematics lessons because Chinese students do not have that much time to directly operate DGS. Therefore, Mrs. Y made a drastic change to the lesson; she let some of her students operate DGS to solve the task.

As shown in the table, we found that Mrs. Y posed some questions focusing on mathematics: "Which one is the center of the circle? Which is the intersecting point?" DGS was not important in these questions (there were no terms related to the software), and the students did not need to use DGS to find the correct answers.

Unlike the other two teachers in our study, we noticed that Mrs. Y posed some feedback that focused on teaching mathematics using technology by using terminology on how to operate DGS, such as "click it", "delete", and "tap". At this time, Mrs. Y wanted the students to know how to use DGS (Figure 5), For her, DGS was a tool that could help the students to solve mathematics tasks.

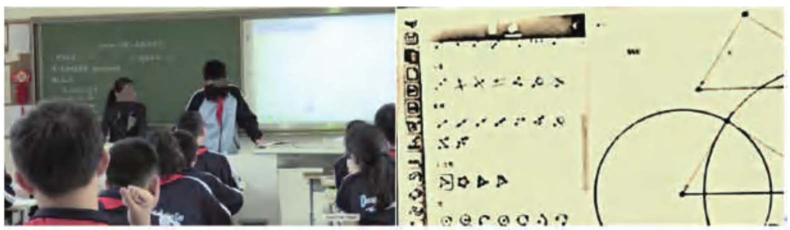


Figure 5. Mrs. Y tried to ask one student to use DGS to construct a triangle.

Based on her teaching experiences, Mrs. Y found that Chinese teachers always told students the answers directly: "I think maybe Chinese teachers will not let the students think by themselves." This may mean that students may not understand the construction method. Therefore, Mrs. Y wanted her students to operate DGS directly and explore the differences between DGS and the use of a pencil and paper. In this way, DGS can help students to develop a clear understanding of geometry.

6. Discussion and Conclusions

This study found that teacher–student interactions are often guided by the initiation–response–evaluation pattern [29]. Additionally, questions and answers may not be derived from students' thought processes during lessons [4]. This implies that teachers may be worried that their students are not familiar with situations in which technology can be used or that they engage in off-task activities, which are discouraged by teachers [30].

Furthermore, although researchers have proposed that technology can create opportunities for students to effectively learn mathematics [4], the teachers in this study seldom linked their questions with DGS. In this study, using DGS did not mean that the teachers changed their interactions with the students. Most of the questions focused only on mathematics, which means that the students were able to answer them without DGS. Therefore, there were not enough opportunities for the students to explore mathematics with the help of DGS in the classroom. This implies that the teachers did not believe that DGS needed to be included in the problem-solving processes of the students. Additionally, they wanted the students to focus on the answers rather than how to explain them using DGS or to consider the concepts of the mathematical tasks [31]. They used DGS as a projector or an amplifier [8,9] without dragging or moving the diagrams. Additionally, the teachers designed many step-by-step tasks for the students to use with technology, with fewer opportunities for the students to engage in mathematical reasoning [4]. This means that the teachers' practices did not present DGS to the students to solve problems [14,15] or explore mathematics [18], but they told the students the mathematical concepts directly [16]. The students were more passive during lessons.

According to research, mathematical interactivity in technology-based classrooms causes students to react to their environment in a mathematical sense. In this research, we assumed that the students must observe these reactions and obtain information on their actions using DGS [6]. This means that the students could use DGS to drag the objects on the screen to explore or check the mathematical content. However, in these lessons, DGS was used more for presenting content. This implies that the dragging mode was controlled by the teachers, not the students; therefore, the meaning of the interactivity of DGS is ambiguous, and DGS was used just for clicking buttons or touching objects on the screen [6]. Using DGS in this way does not impact students' problem-solving processes and does not modify student learning situations and teacher practices in geometry. With DGS, teachers are also required to address new mathematical practices [32,33] and pay more attention to how to design suitable teaching and learning activities [34]. However, the results of this research also imply that the teachers may have lacked sufficient knowledge to help the students via technology [3,35]. Another important factor might be that teachers learned mathematics in a totally different environment to that of modern students; therefore, it was difficult for the teachers to evaluate the students' learning process [3]. In this research, although we could find dynamic elements in each of the tasks used by the teachers, such as moving points, the teachers still used DGS to present the final outcomes and not the entire moving process. When designing some special tasks such as the construction of geometric objects, the teachers chose to interact with the students in a new way [3].

These conclusions concerning the three teachers may not be generalized to other Chinese situations. However, our research seems to present a detailed picture, showing the teachers' strategies regarding the use of DGS in mathematics education. In this regard, further research is required to advance our understanding of how technology resources such as DGS impact the teaching of mathematics. It would be essential to find out whether there are other factors such as a limited class time, heavily didactical tasks, and examinations that may also affect teachers' practices and interactions with students.

Author Contributions: F.Z. conceived, designed, coded, and analyzed the data and drafted the initial manuscript. B.X. supervised and finalized the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding authors.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. An Example of Analyzing Interview Data

Transcription:

Interviewer: what changes do you think DGS take in your lessons and how they can help you during your teaching.

Mrs. J: Of course, it helps me a lot (1) I think technology helps us (2). Now I have used technology for more than 20 years (3). I think I am more and more confident with them (4). I think technology is a supporter (5).

Step 1: We divided this transcription into the following sentences:

Of course, it helps me a lot (1).
I think DGS can help us (2).
Now I have used DGS for more than 20 years (3).
I think I am more and more confident with them (4).
I think technology is a supporter (5).

Step 2: For each sentence, we determined the keywords and their meanings. From the transcript, the first and second sentence had the same meanings. Mrs. J thought that DGS could help her lessons a lot. The third sentence implies that Mrs. J was familiar with DGS. The fourth one shows that Mrs. J was confident when using DGS in the lessons. In addition, the last sentence means that for Mrs. J, DGS was more like a supporter.

Step 3: Afterwards, we discussed each sentence to find out which of them could help us explain how and why the teachers used DGS and interacted with students. For example, in this transcription, because Mrs. J saw DGS as a supporter, she did not always let her students operate DGS, and her questions or feedback always focused on mathematics.

References

1. Martino, A.M.; Maher, C.A. Teacher Questioning to Promote Justification and Generalization in Mathematics: What Research Practice Has Taught Us. *J. Math. Behav.* **1999**, *18*, 53–78. [CrossRef]
2. Bloom, B.S.; Engelhart, M.D.; Furst, E.J.; Hill, W.H.; Krathwohl, D.R. *Taxonomy of Educational Objectives—Handbook 1: Cognitive Domain*; David McKay: New York, NY, USA, 1956.
3. Hollebrands, K.F.; Lee, H.S. Characterizing questions and their focus when pre-service teachers implement dynamic geometry tasks. *J. Math. Behav.* **2016**, *43*, 148–164. [CrossRef]
4. Cayton, C.; Hollebrands, K.; Okumuş, S.; Boehm, E. Pivotal teaching moments in technology-intensive secondary geometry classrooms. *J. Math. Teach. Educ.* **2017**, *20*, 75–100. [CrossRef]
5. Stein, M.K.; Smith, M.S.; Henningsen, M.A.; Silver, E.A. *Implementing Standards-Based Mathematics Instruction: A Casebook for Professional Development*, 2nd ed.; Teachers College Press: New York, NY, USA, 2009.
6. Laborde, C. Interactivity and flexibility exemplified with Cabri. In Proceedings of the Plenary Lecture Given at the Asian Technology Conference in Mathematics, Yogyakarta, Indonesia, 26–30 November 2014; p. 11.
7. Olivero, F. The Proving Process within a Dynamic Geometry Environment. Ph.D. Thesis, Graduate School of Education, Bristol, UK, 2002.
8. Laborde, C. Integration of Technology in the Design of Geometry Tasks with Cabri-Geometry. *Int. J. Comput. Math. Learn.* **2001**, *6*, 283–317. [CrossRef]
9. Soury-Lavergne, S. Duos D'artefacts Tangibles et Numériques et Objets Connectés Pour Apprendre et Faire Apprendre les Mathématiques. Ph.D. Thesis, Ecole Normale Supérieure de Lyon-ENS LYON, Institut Français de l'Education, Lyon, France, 2017.
10. Tall, D. Using the Computer as an Environment for Building and Testing Mathematical Concepts: A Tribute to Richard Skemp. *Paper in Honour Richard Skemp (S. 21–36)*; Mathematics Education Research Centre, University of Warwick, 1986. Available online: http://homepages.warwick.ac.uk/staff/David.Tall/pdfs/dot1986h-computer-skemp.pdf (accessed on 13 August 2014).
11. Healy, L. Identifying and explaining geometrical relationship: Interactions with robust and soft Cabri constructions. In Proceedings of the Conference of the International Group for the Psychology of Mathematics Education (PME), Hiroshima, Japan, 23–27 July 2000; Volume 1, pp. 1–103.
12. Mariotti, M.A. Introduction to proof: The mediation of a dynamic software environment. *Educ. Stud. Math.* **2000**, *44*, 25–53. [CrossRef]
13. Ruthven, K. Expanding Current Practice in Using Dynamic Geometry to Teach about Angle Properties. *Micromath* **2005**, *21*, 26–30.
14. Healy, L.; Hoyles, C. Software Tools for Geometrical Problem Solving: Potentials and Pitfalls. *Int. J. Comput. Math. Learn.* **2001**, *6*, 235–256. [CrossRef]
15. Holzl, R. How does 'dragging' affect the learning of geometry. *Int. J. Comput. Math. Learn.* **1996**, *1*, 169–187. [CrossRef]
16. Guo, J. Let CAI add Wings to Classroom Teaching: On the Characteristics of Geometer's Sketchpad and Its Application in High School Mathematics Teaching. *Math. Teach. Commun.* **2012**, *27*, 16–18.
17. Fan, W. A Case Study of Using Geometric Sketchpad to Carry Out Exploratory Mathematics Learning. *China Educ. Technol.* **2003**, *4*, 34–36.
18. Hu, W. Research of Mathematical Inquiry Teaching Module Based on the Geometer's Sketchpad. Master's Thesis, Guangxi Normal University, Guilin, China, 2005.
19. Liu, H. Integration of Information Technology and Mathematics Curriculum—Geometry Sketchpad—What Can Geometry Sketchpad Do for High School Mathematics Teaching. *Mon. J. High Sch. Math.* **2009**, *9*, 36–38.
20. Pea, R.D. Cognitive Technologies for Mathematics Education. In *Cognitive Science and Mathematics Education*; Erlbaum: Hillsdale, NJ, USA, 1987; pp. 89–122.
21. Özen, D.; Köse, N.Y. Investigating Pre-service Mathematics Teachers' Geometric Problem Solving Process in Dynamic Geometry Environment. *Turk. Online J. Qual. Inq.* **2013**, *4*, 61–74.
22. Koyuncu, I.; Akyuz, D.; Cakiroglu, E. Investigating plane geometry problem-solving strategies of prospective mathematics teachers in technology and paper-and-pencil environments. *Int. J. Sci. Math. Educ.* **2015**, *13*, 837–862. [CrossRef]
23. Brousseau, G. *Theory of Didactical Situations in Mathematics*; Kluwer Academic Publishers: New York, NY, USA, 1997; ISBN 978-0-306-47211-4.
24. Steinbring, H. *The Construction of New Mathematical Knowledge in Classroom Interaction*; Mathematics Education Library; Springer Science & Business Media: Boston, MA, USA, 2005; Volume 38, ISBN 978-0-387-24251-4.
25. Assude, T. Teacher's practices and degree of ICT integration. In Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education, Larnaca, Cyprus, 22–26 February 2007; pp. 1339–1348.
26. Erfjord, I. Teachers' initial orchestration of students' dynamic geometry software use: Consequences for students' opportunities to learn mathematics. *Technol. Knowl. Learn.* **2011**, *16*, 35–54. [CrossRef]

27. Kynigos, C. Designing Constructionist New Mediations for Creative Mathematical Thinking? *Constr. Found.* **2015**, *10*, 305–313.
28. Sabra, H. Contribution à L'étude du Travail Documentaire des Enseignants de Mathématiques: Les Incidents Comme Révélateurs des Rapports Entre Documentations Individuelle et Communautaire. Ph.D. Thesis, Université Claude Bernard-Lyon I, Lyon, France, 2011.
29. Drijvers, P.; Doorman, M.; Boon, P.; Reed, H.; Gravemeijer, K. The teacher and the tool: Instrumental orchestrations in the technology-rich mathematics classroom. *Educ. Stud. Math.* **2010**, *75*, 213–234. [CrossRef]
30. Zbiek, R.M.; Hollebrands, K. A research-informed view of the process of incorporating mathematics technology into classroom practice by inservice and prospective teacher. In *Handbook of Research on Technology in the Learning and Teaching of Mathematics: Syntheses and Perspectives*; Heid, M.K., Blume, G., Eds.; Information Age: Greenwich, CT, USA, 2008; pp. 287–344.
31. Heid, M.K.; Blume, G.W.; Zbiek, R.M.; Edwards, S. Factors that influence teachers learning to do interviews to understand students' mathematical understandings. *Educ. Stud. Math.* **1998**, *37*, 223–249. [CrossRef]
32. Guerrero, S. Technological Pedagogical Content Knowledge in the Mathematics Classroom. *J. Comput. Teach. Educ.* **2010**, *26*, 132–139.
33. Stoilescu, D. Studying Challenges in Integrating Technology in Secondary Mathematics with Technological Pedagogical and Content Knowledge (TPACK). In Proceedings of the International Conferences on Educational Technologies 2014 and Sustainability, Technology and Education 2014, New Tapei City, Taiwan, 10–12 December 2014; pp. 59–66.
34. Hoyles, C.; Lagrange, J.-B. (Eds.) *Mathematics Education and Technology: Rethinking the Terrain: The 17th ICMI Study*; New ICMI Study Series; Springer: New York, NY, USA, 2010; ISBN 978-1-4419-0145-3.
35. Niess, M.L.; Ronau, R.N.; Shafer, K.G.; Driskell, S.O.; Harper, S.R.; Johnston, C.; Browning, C.; Özgün-Koca, S.A.; Kersaint, G. Mathematics Teacher TPACK Standards and Development Model. *Contemp. Issues Technol. Teach. Educ.* **2009**, *9*, 4–24.

sustainability

Review

The Effectiveness of Educational Robots in Improving Learning Outcomes: A Meta-Analysis

Kai Wang [1,†], Guo-Yuan Sang [1,†], Lan-Zi Huang [1,2,*], Shi-Hua Li [3] and Jian-Wen Guo [1]

1 Center for Teacher Education Research, Faculty of Education, Beijing Normal University, Beijing 100875, China
2 School of Education, Hunan First Normal University, Changsha 410205, China
3 College English Teaching Department, Xidian University, Xi'an 710126, China
* Correspondence: 11132021080@bnu.edu.cn
† These authors contributed equally to this work.

Abstract: Numerous studies have been conducted to investigate the potential effect of educational robots, but what appears to be missing is an up-to-date and thorough review of the learning effectiveness of educational robots and the various influencing factors. In this study, a meta-analysis was conducted to systematically synthesize studies' findings on the effects of educational robots on students' learning outcomes. After searching for randomized studies describing educational robots interventions to improve learning outcomes, 34 effect sizes described in 17 articles met the selection criteria. The results of our work evidence a moderate but significantly positive effect of educational robots on learning outcomes ($g = 0.57$, 95% CI [0.49, 0.65], $p < 0.00001$). Moreover, moderator analyses were conducted to investigate important factors relating to the variation of the impact, including educational level and assessment type. Based on the findings of this study, we provide researchers and practitioners with insights into what characteristics of educational robot interventions appear to benefit students' learning outcomes and how pedagogical approaches can be applied in various educational settings to guide the design of future educational robot interventions.

Keywords: educational robots; effectiveness; learning outcomes; meta-analysis

Citation: Wang, K.; Sang, G.-Y.; Huang, L.-Z.; Li, S.-H.; Guo, J.-W. The Effectiveness of Educational Robots in Improving Learning Outcomes: A Meta-Analysis. *Sustainability* **2023**, *15*, 4637. https://doi.org/10.3390/su15054637

Academic Editors: Savvas A. Chatzichristofis, Jian-Hong Ye, Yung-Wei Hao and Yu-Feng Wu

Received: 9 February 2023
Revised: 26 February 2023
Accepted: 27 February 2023
Published: 5 March 2023

1. Introduction

Technological advancements have fundamentally transformed how people, society, and environments inter-relate. Mobilizing digital technology, such as robotics, could significantly facilitate the achievement of the Sustainable Development Goals (SDGs) [1]. As one of the great creative inventions in the 20th century, robots are playing an increasingly important role in industrial intelligent manufacturing, mass production, and public services. Robotics are likely to alter how the SDG are achieved, through replacing and supporting human activities and fostering innovation [2]. In the 1970s, the first educational robot was created in an artificial intelligence laboratory at the Massachusetts Institute of Technology [3]. Early research on educational robots focused on the educational functions of robot kits, including simple kits designed for the purpose of teaching a single function (such as response to sound) as well as complex ones such as Lego Mindstroms that allowed users to build and program [4]. In general, robotic kits are computer-programmed automated machines that are able to perform a series of actions [5]. As more and more social assistance robots (SARs)/social assistance humanoid robots (SAHRs) become available, users will be able to interact with them using actions such as gestures, voice recognition, and emotional expression. SARs are treated as pet animals, toys, or human beings in the form of robots. Given the "uncanny valley" problem [6]. SAHRs are robots with non-threatening appearance, such as Pepper and its precursor, NAO. Because of the ability to talk and show facial expressions, these robots are able to participate in social interactions. For example, they can be used to teach language courses and can even interact with students [7]. Studies concerning the appearance of education robots have examined the user's perception and the physical attributes of the robot (e.g., facial features) [8,9].

With the continuous improvement of robotics technology, educational robots have received great attention from the educational community globally. For example, robotic tutors with empathy have been used to assist elementary school students with learning tasks [10,11]. In South Australia, two schools introduced NAO robots developed and produced by Aldebaran Robotics in France to assist teachers and students [12]. The motivation behind these efforts is that robots can be used to address a variety of challenges faced with education including teacher shortage [13] and teacher workload [14].

The application of educational robots is consistent with several contemporary learning theories such as principles of active learning [15], social constructivism [16], and Papert's constructionism theory [17]. Some evidence is available that the use of robotics in education has a positive impact on student behavior and development, especially in problem-solving skills [18], collaboration [19], learning motivation [20], participation [21], and enjoyment and engagement in the classroom [22,23].These studies drew mixed conclusions about the effectiveness of robotics in education.

Researchers have been actively exploring the use of educational robots in a wide range of courses [24,25]. For example, Hong and colleagues [26] reported that the use of educational robots was beneficial for English learning. Similarly, Toh et al. [27] found that the use of robots helped improve the knowledge of mathematical concepts. McDonald and Howell [28] showed that educational robots could enhance students' interest in engineering and help them gain a better understanding of scientific processes. More broadly, Mathers et al. [29] reported a study that the use of robots enhances knowledge of physics-related topics.

Furthermore, review of literature shows that educational robots are a constantly evolving field with the potential to be implemented in education at all levels from kindergarten to university. Chin et al. [30] indicated that educational robots can provide primary school teachers with tools to increase student achievement. Chang et al. [31] regarded the educational robot as a tool to assist elementary school language teachers. Specifically, educational robots (e.g., NAO) can assist staff in kindergarten by going through a nine-phase procedure [32]. Moreover, NAO and Robovie have also been used to teach children language [33]. Benitti's [34] study reported that the Lego robotics kit is recommended for children age 7 and up. Similarly, Nugent et al. [35] stated that educational robots can teach middle-school students robot activities related to science and engineering processes by giving relatively specific guidance.

Previous systematic review studies have reported the potential contribution of educational robots in schools (e.g., Benitti [34]; Papadopoulos et al., [5]; Spolaôr and Woo et al. [36]; Woo et al. [37]). However, there is a growing criticism from the robotics community in recent years over the lack of empirical research on how robotics can be employed to improve student academic performance [5]. In an earlier study, Benitti's [34] result suggests that few of the empirical studies reviewed support the significance of using educational robots in classroom. Likewise, Woo et al. [37] systematically reviewed studies exploring the possibility of using social robots in naturalistic school settings and identified multiple technical and procedural problems that might affect the successful implementation of such tools. Yet, when examining the overall effectiveness and parameters of successful intervention aimed at the use of educational robots in schools, the above studies are not without limits. For this reason, a meta-analysis was conducted to explore the effectiveness of educational robots in formal learning environments in order to inform, motivate, and guide the use of such tools in future projects. In particular, the study aims to answer the following questions:

Q1. Does the use of educational robots in the classroom improve student learning outcomes?

Q2. Does the effect vary by

- (a) The educational level (pre-school, primary school, secondary school, higher education)?

(b) The subject area (social science and humanities, science)?
(c) The treatment duration (0–4 weeks, 4–8 weeks, above 8 weeks)?
(d) The type of assessment (exam mark, skill-based measure, attitude)?
(e) The robotic type (robotic kits, zoomorphic social robot, humanoid robot)?

2. Method

2.1. Literature Search and Inclusion Criteria

The procedure for selecting studies was based on the Preferred Reporting Items for Systematic reviews and Meta-Analysis (PRISMA) statement [38]. Our search process included two parts. We first consulted the databases of Web of Science and Scopus by using the following search parameters in titles and abstracts of the documents from 2005 to 2023: (robot* OR educational robot*) AND (learning outcome* OR learning achievement* OR academic performance*) AND (student* OR children* OR learner*). Then, we manually screened Google Scholar and additional records identified through citation checking. A study that qualifies for inclusion must examine the use of educational robots and meet the following additional criteria: (1) investigate the effect of an educational robot on student learning; (2) adopt a randomized experimental or quasi-experimental design; (3) include a control group; (4) provide sufficient statistical information for calculating the effect size; (5) be published in a peer-reviewed English language journal; (6) use courses enrolled in a kindergarten, primary school, secondary school, and university or college; and (7) be conducted in a natural school setting.

2.2. Coding Procedure

All records were uploaded to Mendeley. Two researchers independently coded the studies and the following information was extracted for both the experimental group and the control group: the descriptive statistics (i.e., mean, standard deviation, and sample size) of each outcome, the courses involved, sample size, educational levels, robot types, type of assessment, and length of intervention. The inter-coder agreement was evaluated based on both Cohen's Kappa [39] and Gwet's benchmark [40]. All differences in codes were discussed until a consensus was reached. Data were standardized before it was analyzed. First, if studies provide only the p-values, sample size and means, Borenstein et al.'s [41] methods were adopted to estimate standard deviations (*SD*). Second, in studies where pre-test and post-test means were reported, data were combined to generate aggregated means.

When more than one effect size was reported in a study, dependent effect sizes in the meta-analysis were processed according to the characteristic of its dependency to avoid misestimation of standard errors. Various solutions have been proposed in the meta-analysis literature to deal with effect size dependency (e.g., Hedges et al. [42]). For example, one possible solution might be randomly choosing an effect size or taking a mean effect size. However, one problem with this method is that it could result in the loss of data and statistical power [43]. Another problem is that these outcome measures are conceptually different and they may be statistically unrelated [44]. Therefore, if multiple post-test results are reported according to different dimensions in a study, each effect size is considered individually. If several independent sample groups are used, the effect size of each sample group is separately included. In a similar way, in studies involving comparing multiple controls against a single experimental group, the estimated effect size for each pair was separately included. Finally, studies were subgrouped to investigate the possible influence of moderating variables (e.g., the length of intervention). Following previous research, random effects models were applied to investigate the variability of the results of different studies [41].

2.3. Quality Assessment

After preliminary screening of the obtained documents and elimination of duplicate documents and documents that do not match the course matter, the quality of these documents was also examined. The six general sources of bias proposed by Higgins et al. [45] were adopted in the quality assessment including adequacy of allocation sequence generation, allocation concealment, blinding procedures, incomplete result data, selective result reports, and other sources of bias. These items were obtained from published reports, and if more information was needed, we contacted the author. The methodological quality of these items were also checked to determine the level of risk of bias. Low risk of bias is assigned if specious bias might change the outcome; unclear risk of bias is assigned if some suspicion arises; high-risk bias is assigned if paradoxical bias severely affected confidence in the result.

2.4. Statistical Analysis

First, we obtained effect size (*ES*) data from studies included in the meta-analysis. For studies with means and *SD*s, *ES*s were estimated using Review Manager 5.3 [46]. Following previous research (e.g., Tutal and Yazar [47]), we chose to incorporate all *ES*s into the meta-analysis separately. The standardized mean difference index proposed by Hedges [48] was adopted to calculate *ES*s: Hedges = $M_1 - M_2/SD_{\text{pooled}}$, where M_1 refers to the mean score of the treatment group, M_2 refers to the mean score of the control group, and SD_{pooled} refers to the weighted average of the *SD* value of both groups. *ES*s were interpreted based on Cohen's *d*, where 0.8 represents a large effect, 0.5 a medium effect, and 0.2 a small effect. A positive *ES* suggests that the experimental group outperformed the control group. Since continuous data from different scales were extracted, the standardized mean difference (SMD) of the effect size was calculated based on the sample size and 95% confidence interval of each study, and the summary study used analysis of variance. A significance level of 0.05 was set for all analyses (two tailed).

Following the suggestion of Borenstein et al. [41], in case of no heterogeneity, a fixed-effect model was chosen to calculate mean effect size; otherwise a random-effect model was used.To determine what part, if any, of the observed variation was real [41], the I^2 index was used to measure potential heterogeneity [49]. The I^2 value of 25% indicates low level of heterogeneity, 50% indicates moderate level, and 75% indicates high level [50].

2.5. Sensitivity Analysis and Moderator Analyses

The robustness of the results was examined using the leave-one-out method. That is, if the removal of an individual study results in substantial changes, this is an indication of poor homogeneity and therefore the results are unreliable [51]. In meta-analysis, heterogeneity often exists between studies. When multiple moderators are present, they may amplify or attenuate each other's influence on the treatment effectiveness. Hence, moderator analyses were conducted to assess heterogeneity by comparing study subsets [52]. As discussed in the literature review section, several variables could potentially affect the effectiveness of educational robots on student learning (e.g., educational level, discipline, and robotic type).

3. Results

3.1. Search Results

As shown in Figure 1, the initial searches yielded 826 relevant articles. The number was reduced to 269 after duplicates were removed. After examining the title and abstract, another 223 residual references were removed. Full texts were retrieved for 46 articles. In total, 17 articles were retained for further analysis.

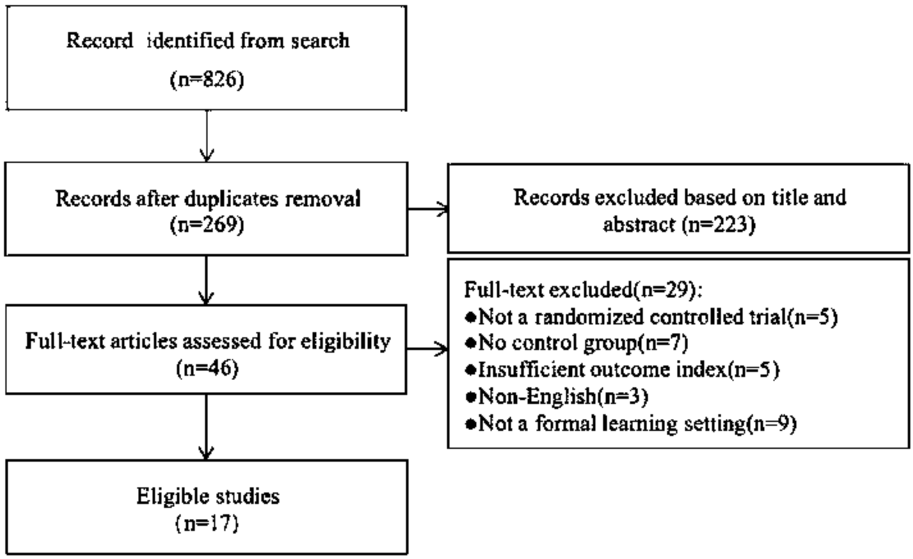


Figure 1. Flow diagram of included studies.

3.2. Characteristics of Included Studies

Table 1 overviews the studies included in the meta-analysis. In one article [53], two studies were identified for inclusion and treated as separate studies. One article [54] had two independent control groups and one experimental group with multiple outcomes, and thus four effect sizes were computed. Four articles reported learning outcomes according to different language skills including listening, speaking, reading, and writing, and thus the effect size for each dimension was treated separately. Final coding led to the inclusion of 34 (k = 34) independent effect sizes from 17 articles.

Table 1. Characteristics of the intervention for each study included in the review.

Study (Year)	Sample Size (E/C)	Discipline	Educational Level	Treatment Duration	Assessment	Robotic Type
Ajlouni (2023) [55]	25/25	Science	Primary education	8 weeks	Intrinsic motivation	LEGO WeDo 2.0 robotic
Alemi et al. (2015) [56]	30/16	English	Secondary education	5 weeks	Anxiety scores	Humanoid robot
Al Hakim et al. (2020) [57]	24/26	Theater	Secondary education	6 weeks	Official drama performance	Social robot
Casad and Jawaharlal (2012) [53]	174/86	Robotics program	Primary education	25 weeks	General academic performance	STEM robotic kits
	65/66	Robotics program	Primary education	6 months	Attitudes toward math	STEM robotic kits
Chen et al. (2013) [58]	30/30	English	Primary education	50 min	Learning achievement	Social robot
Han et al. (2008) [54]	30/30	English	Primary education	40 min	Post-test only achievement	IROBI
	30/30	English	Primary education	40 min	Interest	IROBI

Table 1. *Cont.*

Study (Year)	Sample Size (E/C)	Discipline	Educational Level	Treatment Duration	Assessment	Robotic Type
	30/30	English	Primary education	40 min	Learning achievement	IROBI
	30/30	English	Primary education	40 min	Interest	IROBI
Hong et al. (2016) [26]	25/27	English	Primary education	2 h	Listening	Humanoid robot
	25/27	English	Primary education	2 h	Speaking	Humanoid robot
	25/27	English	Primary education	2 h	Reading	Humanoid robot
	25/27	English	Primary education	2 h	Writing	Humanoid robot
	25/27	English	Primary education	2 h	Learning motivation	Humanoid robot
Hsiao et al. (2015) [59]	30/27	Chinese	Pre-school education	8 weeks	Reading literacy	Social robot iRobiQ
Hsieh et al. (2022) [60]	35/35	Computer concepts	Higher education	8 weeks	Computational thinking capabilities	Humanoid robot
Hyun et al. (2008) [61]	17/17	Korea linguistic ability	Pre-school education	7 weeks	Story making	Social robot iRobiQ
	17/17	Korea linguistic ability	Pre-school education	7 weeks	Story understanding	Social robot iRobiQ
	17/17	Korea linguistic ability	Pre-school education	7 weeks	Vocabulary	Social robot iRobiQ
	17/17	Korea linguistic ability	Pre-school education	7 weeks	Word recognition	Social robot iRobiQ
Julià and Antolí (2016) [62]	9/12	Mathematics	Primary education	8 weeks	Spatial ability average scores	Lego
Korkmaz (2016) [63]	27/26	Computer programming	Higher education	8 weeks	Academic achievement test	Lego Mindstorms Ev3
La Paglia et al. (2011) [64]	15/15	Mathematics	Secondary education	10 weeks	Metacognitive control	Robotic kits
Lindh and Holgersson (2007) [65]	170/161	Programmable construction	Primary education	12 months	Mathematical problems	Lego
	184/160	Programmable construction	Primary education	12 months	Logical problems	Lego

Table 1. *Cont.*

Study (Year)	Sample Size (E/C)	Discipline	Educational Level	Treatment Duration	Assessment	Robotic Type
Ortiz et al. (2017) [66]	33/27	Computer programming	Higher education	16 weeks	The structure of the vehicle and its components	Robotic kits
Wu et al. (2015) [67]	31/33	English	Primary education	4 lecture hours	Learning outcomes	Humanoid robot
	31/33	English	Primary education	4 lecture hours	Learning motivation and interest	Humanoid robot
Yang et al. (2023) [68]	41/34	Information management	Higher education	5 weeks	Academic achievement	AR Bot
	41/34	Information management	Higher education	5 weeks	Enjoyment	AR Bot
	41/34	Information management	Higher education	5 weeks	Problem decomposition skill	AR Bot
	41/34	Information management	Higher education	5 weeks	Algorithm design skill	AR Bot
	41/34	Information management	Higher education	5 weeks	Algorithm efficiency skill	AR Bot

Studies examined the learning effectiveness of educational robots in different courses of two broad categories. Nine articles investigated the use of educational robots in science, technology, engineering, and math (STEM) courses (e.g., C programming) and eight articles analyzed social science courses (e.g., English). Eight articles involved primary children, and three studies had secondary school participants. Two studies had pre-school children and four studies involved students in higher education respectively. The length of intervention differed greatly, ranging from less than 1 h to 25 weeks. The selection included 11 effect sizes for theoretical examination scores, 16 effect sizes for skill examination scores, and 7 effect sizes for attitude towards the course. Seven articles measured the learning effectiveness of robotic kits, five articles examined the learning effectiveness of humanoid robots, and another five articles focused on the learning effectiveness of social robots.

3.3. Study Quality

The results of the risk bias assessment in Figure 2 show that most of the included studies obtain satisfactory scores on the six areas, indicating a low risk of bias. Most studies mentioned that a cluster randomized sampling was used or simply stated that "randomization" was used. Selective reporting bias was also examined. Except for one study with missing data, all the other studies fully reported study results.

3.4. Random-Effect Model Meta-Analysis

3.4.1. Main Effect

The primary goal of this study is to understand the nature of the effect of educational robots on student learning. Figure 3 displays a forest plot of the included studies. The forest plot shows the 95% CIs of the *ES*s of individual studies. Given a high heterogeneity (I^2 = 82%, $p < 0.00001$), a random-effects model was applied. A significant overall effect size ($g = 0.57$, 95% CI [0.49, 0.65], $p < 0.00001$) indicates that teaching methods incorporating educational robots are conducive to learning outcomes. In general, teaching methods using educational robots can improve learning outcomes by 0.57 *SD*, a moderate but significantly positive effect according to Cohen and Lee [69]. In addition, the sensitivity analysis showed the results were robust, and the effect size did not vary considerably when the leave-one-out method was used. Regarding sensitivity analysis, the effect size did not vary considerably neither when results were computed by using the leave-one-out method.

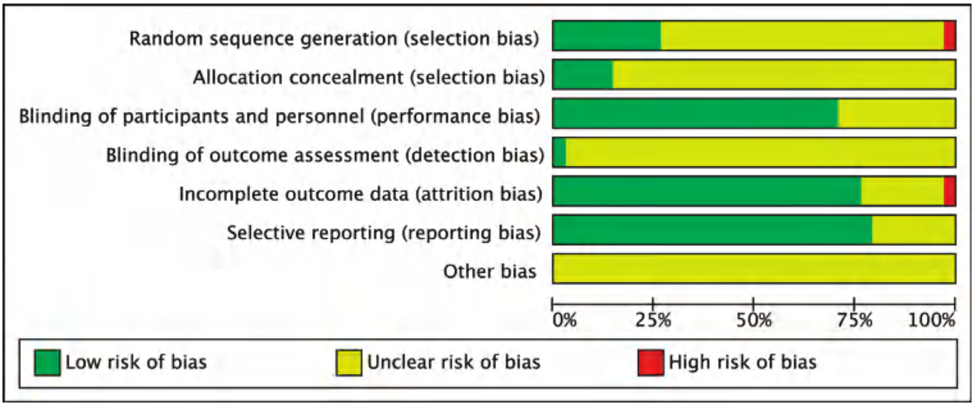
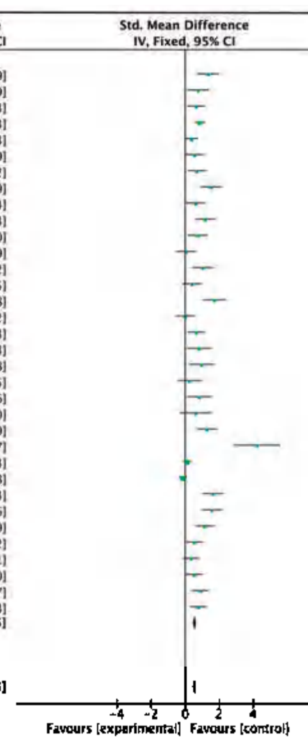


Figure 2. Risk of bias.



Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
1.1.5 all reference								
Ajlouni (2023)	3.2	0.24	25	2.86	0.25	25	1.6%	1.37 [0.75, 1.99]
Alemi et al. (2015)	3.48	0.52	30	3	0.77	16	1.6%	0.76 [0.14, 1.39]
Al Hakim et al. (2020)	207.46	28.67	35	189.6	26.1	35	2.7%	0.64 [0.16, 1.13]
Casad & Jawaharlal (2012.1)	8.47	1.83	174	6.95	1.63	86	8.6%	0.86 [0.59, 1.13]
Casad & Jawaharlal (2012.2)	4.82	0.86	65	4.48	0.9	66	5.2%	0.38 [0.04, 0.73]
Chen et al. (2013)	80.83	23.33	30	68	20.81	30	2.3%	0.57 [0.06, 1.09]
Han et al. (2008.1)	5.5	1.13	30	4.61	1.37	30	2.3%	0.70 [0.18, 1.22]
Han et al. (2008.2)	4.47	0.68	30	3.4	0.72	30	1.9%	1.51 [0.93, 2.09]
Han et al. (2008.3)	5.5	1.13	30	4.8	1.09	30	2.3%	0.62 [0.10, 1.14]
Han et al. (2008.4)	4.47	0.68	30	3.57	0.82	30	2.0%	1.18 [0.63, 1.73]
Hong et al. (2016.1)	36.22	5.76	25	31.11	7.71	27	2.0%	0.74 [0.17, 1.30]
Hong et al. (2016.2)	11.2	1.63	25	11.11	2.03	27	2.1%	0.05 [−0.50, 0.59]
Hong et al. (2016.3)	26.4	4.66	25	20.56	6.27	27	1.8%	1.04 [0.45, 1.62]
Hong et al. (2016.4)	15.72	3.16	25	14.11	4.62	27	2.1%	0.40 [−0.15, 0.95]
Hong et al. (2016.5)	135.2	6.99	25	107.66	20.65	27	1.5%	1.73 [1.09, 2.38]
Hsiao et al. (2015)	7.77	7.17	30	7.8	4.74	27	2.3%	−0.00 [−0.52, 0.52]
Hsieh et al.(2022)	207.46	28.67	35	189.6	26.1	35	2.7%	0.64 [0.16, 1.13]
Hyun et al. (2008.1)	2.47	0.86	17	1.86	0.54	17	1.3%	0.83 [0.13, 1.53]
Hyun et al. (2008.2)	2.47	1.05	17	1.31	1.28	17	1.2%	0.97 [0.25, 1.68]
Hyun et al. (2008.3)	5.21	0.83	17	5	0.64	17	1.4%	0.28 [−0.40, 0.95]
Hyun et al. (2008.4)	8.35	5.93	17	3.76	4.52	17	1.2%	0.85 [0.14, 1.56]
Julià & Antolí (2016)	7.9	8.95	9	2	9.4	12	0.8%	0.61 [−0.27, 1.50]
Korkmaz (2016)	66.8	16.9	26	45.2	16	27	1.7%	1.29 [0.70, 1.89]
La Paglia et al. (2011)	2.77	2.09	15	−6.3	2.09	15	0.3%	4.22 [2.87, 5.57]
Lindh & Holgersson (2007.1)	29.44	5.04	170	28.84	5.46	161	13.3%	0.11 [−0.10, 0.33]
Lindh & Holgersson (2007.2)	11.16	2.96	184	11.53	2.45	160	13.8%	−0.13 [−0.35, 0.08]
Ortiz et al. (2017)	86.15	8.62	33	70.41	10.5	27	1.8%	1.63 [1.04, 2.23]
Wu et al. (2015.1)	91.06	5.17	31	79.93	8.18	33	1.9%	1.60 [1.03, 2.16]
Wu et al. (2015.2)	4.2	0.76	31	3.3	0.77	33	2.2%	1.16 [0.63, 1.69]
Yang et al. (2023.1)	3.82	0.77	41	3.3	1.07	34	2.9%	0.56 [0.10, 1.02]
Yang et al. (2023.2)	81.57	6.7	41	79.09	7.38	34	3.0%	0.35 [−0.11, 0.81]
Yang et al. (2023.3)	3.51	0.8	41	3.02	1.02	34	2.9%	0.54 [0.07, 1.00]
Yang et al. (2023.4)	3.52	0.79	41	2.68	1.07	34	2.7%	0.90 [0.42, 1.37]
Yang et al. (2023.5)	3.23	0.85	41	2.53	0.98	34	2.8%	0.76 [0.29, 1.23]
Subtotal (95% CI)			**1441**			**1281**	**100.0%**	**0.57 [0.49, 0.65]**
Heterogeneity: Chi² = 180.77, df = 33 (P < 0.00001); I² = 82%								
Test for overall effect: Z = 14.27 (P < 0.00001)								
Total (95% CI)			**1441**			**1281**	**100.0%**	**0.57 [0.49, 0.65]**
Heterogeneity: Chi² = 180.77, df = 33 (P < 0.00001); I² = 82%								
Test for overall effect: Z = 14.27 (P < 0.00001)								
Test for subgroup differences: Not applicable								

Figure 3. Forest plot of the effect sizes [26,53–68].

The funnel plot is a common method for qualitatively measuring publication bias, and it is based on the hypothetical design that the accuracy of the estimation of the effect of intervention measures increases with the increase in the sample size. Figure 4 shows the result of the funnel plot for the 17 articles (k = 34) and it suggests no significant publication bias.

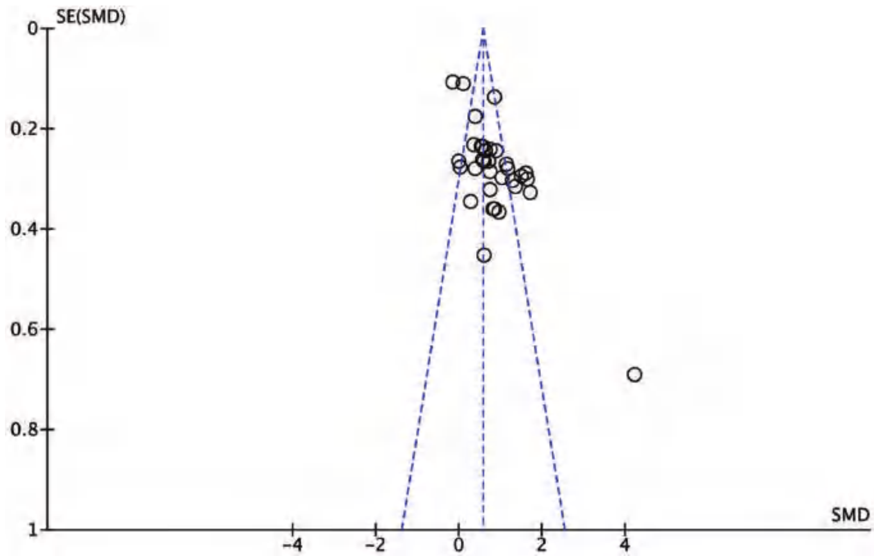


Figure 4. Funnel plot. Note: Each circle represents one effect size included in the meta-analysis.

3.4.2. Moderator Analyses

We conducted moderator analyses to determine the possible explanations for the high heterogeneity among studies. The random-effect mode was chosen to explore the effect of potential moderator variables identified in this study, including educational level, subject area, treatment duration, assessment type, and robotic type.

Table 2 shows that the SMD of each subset was positive and exclusive of zero, indicating that students who used educational robots achieved higher learning effectiveness than those who did not. In addition, educational level and assessment type were significantly related to the variability in the learning effectiveness. However, three moderators were not significant: the subject area, treatment duration, and the robotic type. The following sections present the results for each moderator.

To examine if the influence of educational robot-based classroom instruction varies across educational levels, studies were divided into four subsets based on research setting: pre-school setting, primary school setting, secondary school setting, and higher education setting. As indicated in Table 2, educational robots were found to have positive effects on student learning at all educational levels. In terms of the strength of the effect, educational robots had quite a strong effect on student learning at secondary school level (g = 1.69) and higher education level (g = 1.42), and had a less strong effect at primary school level (g = 0.78) and pre-school level (g = 0.55).

Table 2. Results for the moderator analyses.

Moderator Variables	*k*	*SMD*	*Z*	I^2 (%)	*p*
Educational level				73.5	0.01 *
1. Pre-school	5	0.55	2.64		
2. Primary school	18	0.78	5.13		
3. Secondary school	3	1.69	2.14		
4. Higher education	8	1.42	6.76		
Subject area				0	0.69
1. Social science and humanities	19	0.80	7.05		
2. Science	15	0.87	3.46		
Treatment duration				0	0.57
1. 0–4 weeks	12	0.92	6.17		
2. 4–8 weeks	13	0.72	5.01		
3. Above 8 weeks	9	0.79	3.43		
Type of assessment				83.7	0.046 *
1. Exam mark	11	0.97	4.84		
2. Skill-based measure	16	0.49	3.56		
3. Attitude	7	1.23	8.08		
Robotic type				0	0.54
1. Robotic kits	9	0.88	3.56		
2. Zoomorphic social robot	11	0.71	5.36		
3. Humanoid robot	14	0.91	4.53		

Note: * $p < 0.05$.

The included studies involved a diverse range of courses. To examine whether the use of educational robots in classroom instruction is more beneficial for some courses than for others, we separated the courses into science courses (e.g., mathematics and computer programming) and non-science courses (e.g., English and theater). Eligible studies were equally distributed between the two subsets. We also compared the treatment duration using moderator analysis. The results suggested that subject area (I^2 = 0%, $p > 0.05$) and treatment duration (I^2 = 0%, $p > 0.05$) had no significant moderating influence on the relationship between the use of educational robots and learning outcomes. In other words, the effects of educational robots used in different disciplines were not different, and the effects of implementing robotics-based educational tools from less than an hour to above eight weeks were not significant.

The learning outcomes of educational robot-assisted learning were often assessed using an examination score or some skill-based measure. Further, self-report questionnaires were also used to elicit student attitudes toward the course. It was found that the effects of educational robots in the classroom were totally different according to the type of assessment (I^2 = 83.7%, $p < 0.05$). The results show that the implementation of educational robots in student attitudes (g = 1.23) were significantly better than in exam mark (g = 0.97) or tests of skill (g = 0.49). Further, no significant heterogeneity value was found within each robotic type (I^2 = 0%, $p > 0.05$). The *ES*s produced by studies implementing robotic kits, social robots, and humanoid robots were +0.88, +0.71, and +0.91, respectively.

4. Discussion

4.1. The Learning Effectiveness of Educational Robots

This study meta-analyzed 17 articles that examined the learning effectiveness of educational robots. Educational robots were found to have a moderate but significantly positive effect on student learning outcomes. This finding suggests that educational robot-based classroom instruction tends to produce better learning outcomes than traditional lecture-style teaching. This finding corroborates the positive results reported in previous review studies [5,34,37]. Several possible explanations might contribute to why students who received robot-based classroom instruction had better learning outcomes than those who were taught by traditional teaching methods.

First, some advantageous features of educational robots including the possibility of performing repetitive tasks precisely [31] can be leveraged for teaching purposes if they match the instructional goals. Woo et al. [37] found that social assistance humanoid robots (e.g., Pepper and NAO) have been used to support classroom teaching by reducing time-consuming and tedious repetitive activities. As a teaching tool, the repeatability of educational robots would provide more in-class time for collaborative learning activities. Thus, teachers are allowed more time to monitor student learning progress and provide timely assistance when problems occur.

Another explanation for this finding may be that educational robots can facilitate learning [70]. The use of educational robots can help to inspire curiosity and an enjoyable learning environment via interesting activities and practical experiences [23]. In addition, Alemi et al. [56] argued that the pleasure of the robot brought interesting interactions and lowered students' anxiety during the learning activities. The robot's friendly attitude towards children added an element of non-judgment and comfort to the interaction while learning, which in turn lowered the fear of making mistakes.

Third, previous reviews (e.g., Cheung and Slavin [71]) suggested that a sample size of 250 or fewer can be treated as small. Small sample size is a common concern for nearly all studied in the meta-analysis. Compared with those with large sample sizes, studies with small sample sizes tend to be more strictly controlled and thus there is a higher chance of obtaining positive results [72]. In addition, the file-drawer effect is more likely to occur in studies with small sample sizes where null effects are found.

4.2. Moderators for Educational Robots on Learning Effectiveness

Five factors were included in the moderator analysis. With regard to the factor of the educational level, results showed that the difference between the summated *ESs* of educational levels was found to be statistically significant; however, given the small number of studies included, interpretation of the *ES*s must be carried out with caution [73]. There were quite strong effects for secondary school students (note that $k = 3$) and higher education students (note $k = 8$), moderately strong effect for pre-school children, and strong effect for primary school students. These divergent findings might not be easily interpreted by a discerning pattern. However, one might tentatively speculate that robots are still limited in their capacities to perceive the human world [74]. Moreover, robots are typically used in situations where the lessons are short and well-structured, and delivered with little adaptation to the needs of an individual learner or a curriculum [75]. Therefore, the small effect of educational robots on student learning might be attributed to factors such as the learning content, activity format, or course requirements at the pre-school and primary levels.

In addition, Fernández-Llamas et al. [76] noted that younger students were more familiar with robots and were also more likely to believe that robots could think [77]. However, findings of some studies suggest otherwise. For example, Serholt [78] found that children expected that a robot could understand their intentions like a human teacher does. When the robot failed to do so, children would perform the task on their own or stop interacting with the robot. The study also pointed out that it is paradoxical to see that robots with a humanoid appearance actively participated in a learning activity but they were deficient in social interaction and cooperation skills were non-existent. Together, these findings suggest that the adoption of educational robots at the pre-school and primary levels requires a more careful instructional design. Moreover, looking ahead, we hope that robots someday reach an adequate level of humanlike perception and communication abilities to play such a role in interpreting their intentions, much like human teachers do.

Second, with regard to the type of assessment, the difference between the summated *ES*s of the groups was found to be statistically significant. The meta-analysis revealed that educational robots had a stronger impact on student attitude and were slightly weaker in improving theoretical and skill examination scores. Chin et al.'s [30] study found that students' attitude toward the educational robot was positive. In another study, Benitti [34] reported that the young people with different interests could potentially benefit from educational robot-based instruction. It is possible that interaction with robots can increase student motivation, engagement, and attitude towards education. Given these findings, future research on the application of educational robots should consider exploring how these tools in school environments can be better employed to promote the learning of knowledge and skills.

Third, statistically significant differences were not found among the moderators of subject area and robotics type. This seems to be an unsurprising result. The effectiveness of educational robots depends on various factors, and they must be adaptable in real time [7]. In general, there appears to be no discrepancy in learning outcomes if the robot chosen is suitable for the specific discipline. Prior studies showed that educational robots can help improve student learning outcomes in mathematics and science courses [27,79]. Chang et al. [31] reported that socially assistive robots can promote English language skills. However, to meet the learning purposes of incorporating robots in classroom settings, they must be deployed and studied in the context of their primary use. For example, it is unreasonable to use complex robots with pre-school children in a classroom setting. As such, future research has a responsibility to explore the use of different of robotics in terms of different learning activities.

Finally, the three categories of treatment duration are proved to be insignificant. This finding is in line with results of recent reviews (e.g., Cheung and Slavin [71]) which have consistently found that technology has no effect in improving student learning in the long run. Moreover, another factor for the nonsignificant result might be learning intensity because not every study reviewed in this meta-analysis indicated if students spent an equal amount of time using educational robots.

5. Conclusions

Undoubtedly, educational robots will be expected to take on a more vital role in schools in the future. Therefore, how educational robots can be best integrated into classroom instruction is a question that deserves greater attention. Despite the proposed advantages of incorporating educational robots in student learning [5,23,34,80], currently researchers have not given particular attention to the overall effectiveness and parameters of successful intervention aiming at the use of educational robots in school settings. To address this gap, this study, therefore, conducted a meta-analysis of the effects of educational robots in the classroom.

This study found supporting evidence for the positive effects of educational robot-based interventions on student learning (mean ES = +0.57), suggesting that educational robots can be leveraged to facilitate student learning. Moreover, the homogeneity test indicated that there was a high level of heterogeneity among the effect sizes of the studies reviewed in this meta-analysis. To further investigate this heterogeneity, five factors including course type, education level, treatment duration, assessment type, and robot type were quantitatively assessed using the moderator analysis. Our findings partially support and enrich the existing research in various ways. Some of the findings provide guidance and direction for the process of educational robot operation. In terms of treatment duration, the usefulness of robotics-based instruction remains stable as the duration of implementation is extended. What is clear from investigating the moderator effect of subject area and treatment duration is that the choice of robot usually depends on practical considerations.

Overall, course designers and teachers can use these results in course design and facilitation of learning to improve student's learning in educational robot-based courses as well as differentiating their practices according to course level (e.g., children and undergraduate students), discipline, and robotic type. Finally, we hope that the results of this study can further advance our understanding of implementing educational robots in formal learning settings, and enhance the quality of education for students engaged in robot-based learning.

6. Limitations and Future Research

A first limitation was that this study only focused on several factors that might affect the effects of educational robot interventions. It did not consider other factors such as gender differences or socio-economic status which might also influence student learning outcomes. Future research should consider how these and other factors might be related to students' academic performance involving the use of robotic technology in education settings.

Another limitation was that most randomized-controlled studies included in the meta-analysis were conducted with children aged 8 to 12 years old. Further research should be conducted to explore how these findings can be applied to different age groups.

Finally, although it is central to analyze specific characteristics of the intervention's influence on the effect of educational robots in formal learning environments, in the future it will be valuable to explore interaction effects among moderators, which can provide valuable information to answer questions such as "do these intervention components amplify or attenuate each other's effectiveness?"

Author Contributions: Conceptualization, K.W. and G.-Y.S.; methodology, S.-H.L. and L.-Z.H.; writing—review and editing, S.-H.L., L.-Z.H., K.W. and J.-W.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by "the Fundamental Research Funds for the Central Universities" (2022NTSS17).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not Applicable.

Data Availability Statement: The datasets generated and/or analyzed during the current study are available from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sachs, J.D.; Schmidt-Traub, G.; Mazzucato, M.; Messner, D.; Nakicenovic, N.; Rockström, J. Six transformations to achieve the Sustainable Development Goals. *Nat. Sustain.* **2019**, *2*, 805–814. [CrossRef]
2. Guenat, S.; Purnell, P.; Davies, Z.G.; Nawrath, M.; Stringer, L.C.; Babu, G.R.; Balasubramanian, M.; Ballantyne, E.E.F.; Bylappa, B.K.; Chen, B.; et al. Meeting sustainable development goals via robotics and autonomous systems. *Nat. Commun.* **2022**, *13*, 3559. [CrossRef] [PubMed]
3. Papert, S.; Solomon, C. *Twenty Things to Do with a Computer*; Constructing Modern Knowledge Press: Cambridge, MA, USA, 1971.
4. Karim, M.E.; Lemaignan, S.; Mondada, F. A review: Can robots reshape K-12 STEM education? In Proceedings of the 2015 IEEE International Workshop on Advanced Robotics and Its Social Impacts (ARSO), Lyon, France, 30 June–2 July 2015; pp. 1–8.
5. 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.* **2020**, *155*, 103924. [CrossRef]
6. Li, D.; Rau, P.P.; Li, Y. A cross-cultural study: Effect of robot appearance and task. *Int. J. Soc. Robot.* **2010**, *2*, 175–186. [CrossRef]
7. Mubin, O.; Stevens, C.J.; Shahid, S.; Al Mahmud, A.; Dong, J.J. A review of the applicability of robots in education. *J. Technol. Educ. Learn.* **2013**, *1*, 13. [CrossRef]
8. Barker, B.S.; Ansorge, J. Robotics as means to increase achievement scores in an informal learning environment. *J. Res. Technol. Educ.* **2007**, *39*, 229–243. [CrossRef]
9. Woods, S.; Dautenhahn, K.; Schulz, J. The design space of robots: Investigating children's views. In Proceedings of the 13th IEEE International Workshop on Robot and Human Interactive Communication, Kurashiki, Japan, 22 September 2004; pp. 47–52.
10. Serholt, S.; Barendregt, W. Robots tutoring children: Longitudinal evaluation of social engagement in child-robot interaction. In Proceedings of the 9th Nordic Conference on Human-Computer Interaction (NordiCHI'16), Gothenburg, Sweden, 23–27 October 2016; pp. 1–10.
11. Obaid, M.; Aylett, R.; Barendregt, W.; Basedow, C.; Corrigan, L.J.; Hall, L.; Jones, A.; Kappas, A.; Kuster, D.; Paiva, A.; et al. Endowing a robotic tutor with empathic qualities: Design and pilot evaluation. *Int. J. Hum. Robot.* **2018**, *15*, 1850025. [CrossRef]
12. Keane, T.; Chalmers, C.; Boden, M.; Williams, M. Humanoid robots: Learning a programming language to learn a traditional language. *Technol. Pedagog. Educ.* **2019**, *28*, 533–546. [CrossRef]

13. Edwards, B.I.; Cheok, A.D. Why not robot teachers: Artificial intelligence for addressing teacher shortage. *Appl. Artif. Intell.* **2018**, *32*, 345–360. [CrossRef]
14. Movellan, J.R.; Tanaka, F.; Fortenberry, B.; Aisaka, K. The RUBI/QRIO project: Origins, principles, and first steps. In Proceedings of the 4th IEEE International Conference on Development and Learning (ICDL-2005), Osaka, Japan, 19–21 July 2005; pp. 80–86.
15. Harmin, M.; Toth, M. *Inspiring Active Learning: A Complete Handbook for Today's Teachers*; ASCD: Alexandria, VA, USA, 2006.
16. Vygotsky, L.S. *Mind in Society: The Development of Higher Psychological Processes*; Harvard University Press: Cambridge, MA, USA, 1980.
17. Papert, S. *Mindstorms: Children, Computers and Powerful Ideas*; Basic Books: New York, NY, USA, 1980.
18. Barak, M.; Zadok, Y. Robotics projects and learning concepts in science, technology and problem solving. *Int. J. Technol. Des. Educ.* **2009**, *19*, 289–307. [CrossRef]
19. Hong, J.C.; Yu, K.C.; Chen, M.Y. Collaborative learning in technological project design. *Int. J. Technol. Des. Educ.* **2011**, *21*, 335–347. [CrossRef]
20. Kubilinskienė, S.; Žilinskienė, I.; Dagienė, V.; Sinkevičius, V. Applying robotics in school education: A systematic review. *Balt. J. Mod. Comput.* **2017**, *5*, 50–69. [CrossRef]
21. Rusk, N.; Resnick, M.; Berg, R.; Pezalla-Granlund, M. New pathways into robotics: Strategies for broadening participation. *J. Sci. Educ. Technol.* **2008**, *17*, 59–69. [CrossRef]
22. Han, J.; Kim, D. r-Learning services for elementary school students with a teaching assistant robot. In Proceedings of the 2009 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI), La Jolla, CA, USA, 11–13 March 2009; pp. 255–256.
23. Alimisis, D. Educational robotics: Open questions and new challenges. *Themes Sci. Technol. Educ.* **2013**, *6*, 63–71.
24. Baker, T.; Smith, L.; Anissa, N. Educ-AI-tion Rebooted? Exploring the Future of Artificial Intelligence in Schools and Colleges. NESTA. 2019. Available online: https://www.nesta.org.uk/report/education-rebooted/ (accessed on 8 February 2023).
25. Tuomi, I. *The Impact of Artificial Intelligence on Learning, Teaching, and Education: Policies for the Future*; JRC Science for Policy Report; Publications Office of the European Union: Luxembourg, 2018.
26. Hong, Z.W.; Huang, Y.M.; Hsu, M.; Shen, W.W. Authoring robot-assisted instructional materials for improving learning performance and motivation in EFL classrooms. *J. Educ. Technol. Soc.* **2016**, *19*, 337–349.
27. Toh, L.P.E.; Causo, A.; Tzuo, P.W.; Chen, I.M.; Yeo, S.H. A review on the use of robots in education and young children. *J. Educ. Technol. Soc.* **2016**, *19*, 148–163.
28. McDonald, S.; Howell, J. Watching, creating and achieving: Creative technologies as a conduit for learning in the early years. *Br. J. Educ. Technol.* **2012**, *43*, 641–651. [CrossRef]
29. Mathers, N.; Goktogen, A.; Rankin, J.; Anderson, M. Robotic mission to mars: Hands-on, minds-on, web-based learning. *Acta Astronaut.* **2012**, *80*, 124–131. [CrossRef]
30. Chin, K.Y.; Hong, Z.W.; Chen, Y.L. Impact of using an educational robot-based learning system on students' motivation in elementary education. *IEEE Trans. Learn. Technol.* **2014**, *7*, 333–345. [CrossRef]
31. Chang, C.W.; Lee, J.H.; Chao, P.Y.; Wang, C.Y.; Chen, G.D. Exploring the possibility of using humanoid robots as instructional tools for teaching a second language in primary school. *J. Educ. Technol. Soc.* **2010**, *13*, 13–24.
32. Fridin, M. Kindergarten social assistive robot: First meeting and ethical issues. *Comput. Hum. Behav.* **2014**, *30*, 262–272. [CrossRef]
33. Uluer, P.; Akalın, N.; Köse, H. A new robotic platform for sign language tutoring. *Int. J. Soc. Robot.* **2015**, *7*, 571–585. [CrossRef]
34. Benitti, F.B.V. Exploring the educational potential of robotics in schools: A systematic review. *Comput. Educ.* **2012**, *58*, 978–988. [CrossRef]
35. Nugent, G.; Barker, B.; Grandgenett, N.; Adamchuk, V.I. Impact of robotics and geospatial technology interventions on youth STEM learning and attitudes. *J. Res. Technol. Educ.* **2010**, *42*, 391–408. [CrossRef]
36. Spolaôr, N.; Benitti, F.B.V. Robotics applications grounded in learning theories on tertiary education: A systematic review. *Comput. Educ.* **2017**, *112*, 97–107. [CrossRef]
37. Woo, H.; LeTendre, G.K.; Pham-Shouse, T.; Xiong, Y. The Use of Social Robots in Classrooms: A Review of Field-based Studies. *Educ. Res. Rev.* **2021**, *33*, 100388. [CrossRef]
38. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Ann. Intern. Med.* **2009**, *151*, 264–269. [CrossRef] [PubMed]
39. Hallgren, K.A. Computing inter-rater reliability for observational data: An overview and tutorial. *Tutor. Quant. Methods Psychol.* **2012**, *8*, 23. [CrossRef]
40. Gwet, K.L. Benchmarking inter-rater reliability coefficients. In *Handbook of Inter-Rater Reliability*, 3rd ed.; Advanced Analytics, LLC: Gaithersburg, MD, USA, 2012; pp. 121–128.
41. Borenstein, M.; Cooper, H.; Hedges, L.; Valentine, J. Effect sizes for continuous data. In *The Handbook of Research Synthesis and Meta-Analysis*; Russell Sage Foundation: Manhattan, NY, USA, 2009; Volume 2, pp. 221–235.
42. Hedges, L.V.; Tipton, E.; Johnson, M.C. Robust variance estimation in meta-regression with dependent effect size estimates. *Res. Synth. Methods* **2010**, *1*, 39–65. [CrossRef] [PubMed]
43. Scammacca, N.; Roberts, G.; Stuebing, K.K. Meta-analysis with complex research designs: Dealing with dependence from multiple measures and multiple group comparisons. *Rev. Educ. Res.* **2014**, *84*, 328–364. [CrossRef] [PubMed]
44. Littell, J.H.; Corcoran, J.; Pillai, V. *Systematic Reviews and Meta-Analysis*; Oxford University Press: Oxford, UK, 2008.

45. Higgins, J.P.; Thomas, J.; Chandler, J.; Cumpston, M.; Li, T.; Page, M.J.; Welch, V.A. (Eds.) *Cochrane Handbook for Systematic Reviews of Interventions*; John Wiley & Sons: Hoboken, NJ, USA, 2019.
46. Collaboration, C. *Review Manager (RevMan) Version 5.3*; The Nordic Cochrane Centre: Copenhagen, Denmark, 2014.
47. Tutal, Ö.; Yazar, T. Flipped classroom improves academic achievement, learning retention and attitude towards course: A meta-analysis. *Asia Pac. Educ. Rev.* **2021**, *22*, 655–673. [CrossRef]
48. Hedges, L.V. Fitting categorical models to effect sizes from a series of experiments. *J. Educ. Stat.* **1982**, *7*, 119–137. [CrossRef]
49. Higgins, J.P.; Thompson, S.G. Quantifying heterogeneity in a meta-analysis. *Stat. Med.* **2002**, *21*, 1539–1558. [CrossRef] [PubMed]
50. Higgins, J.P.; Thompson, S.G.; Deeks, J.J.; Altman, D.G. Measuring inconsistency in meta-analyses. *BMJ* **2003**, *327*, 557–560. [CrossRef]
51. Viechtbauer, W.; Cheung MW, L. Outlier and influence diagnostics for meta-analysis. *Res. Synth. Methods* **2010**, *1*, 112–125. [CrossRef]
52. Li, X.; Dusseldorp, E.; Su, X.; Meulman, J.J. Multiple moderator meta-analysis using the R-package Meta-CART. *Behav. Res. Methods* **2020**, *52*, 2657–2673. [CrossRef] [PubMed]
53. Casad, B.J.; Jawaharlal, M. Learning through guided discovery: An engaging approach to K-12 STEM education. In Proceedings of the 2012 ASEE Annual Conference & Exposition, San Antonio, TX, USA, 10–13 June 2012.
54. Han, J.H.; Jo, M.H.; Jones, V.; Jo, J.H. Comparative study on the educational use of home robots for children. *J. Inf. Proc. Syst.* **2008**, *4*, 159–168. [CrossRef]
55. Ajlouni, A. The Impact of Instruction-Based LEGO WeDo 2.0 Robotic and Hypermedia on Students' Intrinsic Motivation to Learn Science. *Int. J. Interact. Mob. Technol.* **2023**, *17*, 22–39. [CrossRef]
56. Alemi, M.; Meghdari, A.; Ghazisaedy, M. The impact of social robotics on L2 learners' anxiety and attitude in English vocabulary acquisition. *Int. J. Soc. Robot.* **2015**, *7*, 523–535. [CrossRef]
57. Al Hakim, V.G.; Yang, S.H.; Tsai, T.H.; Lo, W.H.; Wang, J.H.; Hsu, T.C.; Chen, G.D. Interactive Robot as Classroom Learning Host to Enhance Audience Participation in Digital Learning Theater. In Proceedings of the 2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT), Tartu, Estonia, 6–9 July 2020; pp. 95–97.
58. Chen, G.D.; Nurkhamid; Wang, C.Y.; Yang, S.H.; Lu, W.Y.; Chang, C.K. Digital learning playground: Supporting authentic learning experiences in the classroom. *Interact. Learn. Environ.* **2013**, *21*, 172–183. [CrossRef]
59. Hsiao, H.S.; Chang, C.S.; Lin, C.Y.; Hsu, H.L. "iRobiQ": The influence of bidirectional interaction on kindergarteners' reading motivation, literacy, and behavior. *Interact. Learn. Environ.* **2015**, *23*, 269–292. [CrossRef]
60. Hsieh, M.C.; Pan, H.C.; Hsieh, S.W.; Hsu, M.J.; Chou, S.W. Teaching the concept of computational thinking: A STEM-based program with tangible robots on project-based learning courses. *Front. Psychol.* **2022**, *12*, 6628. [CrossRef] [PubMed]
61. Hyun, E.J.; Kim, S.Y.; Jang, S.; Park, S. Comparative study of effects of language instruction program using intelligence robot and multimedia on linguistic ability of young children. In Proceedings of the RO-MAN 2008-The 17th IEEE International Symposium on Robot and Human Interactive Communication, Munich, Germany, 1–3 August 2008; pp. 187–192.
62. Julià, C.; Antolí, J.Ò. Spatial ability learning through educational robotics. *Int. J. Technol. Des. Educ.* **2016**, *26*, 185–203. [CrossRef]
63. Korkmaz, Ö. The Effect of Lego Mindstorms Ev3 Based Design Activities on Students' Attitudes towards Learning Computer Programming, Self-Efficacy Beliefs and Levels of Academic Achievement. *Balt. J. Mod. Comput.* **2016**, *4*, 994–1007.
64. La Paglia, F.; Rizzo, R.; La Barbera, D. Use of robotics kits for the enhancement of metacognitive skills of mathematics: A possible approach. *Annu. Rev. Cyberther. Telemed.* **2011**, *167*, 26–30.
65. Lindh, J.; Holgersson, T. Does Lego training stimulate pupils' ability to solve logical problems? *Comput. Educ.* **2007**, *49*, 1097–1111. [CrossRef]
66. Ortiz, O.O.; Franco, J.Á.P.; Garau, P.M.A.; Martín, R.H. Innovative mobile robot method: Improving the learning of programming languages in engineering degrees. *IEEE Trans. Educ.* **2016**, *60*, 143–148. [CrossRef]
67. Wu, W.C.V.; Wang, R.J.; Chen, N.S. Instructional design using an in-house built teaching assistant robot to enhance elementary school English-as-a-foreign-language learning. *Interact. Learn. Environ.* **2015**, *23*, 696–714. [CrossRef]
68. Yang, F.C.O.; Lai, H.M.; Wang, Y.W. Effect of augmented reality-based virtual educational robotics on programming students' enjoyment of learning, computational thinking skills, and academic achievement. *Comput. Educ.* **2023**, *195*, 104721. [CrossRef]
69. Cohen, L.; Lee, C. Instantaneous frequency, its standard deviation and multicomponent signals. In *Advanced Algorithms and Architectures for Signal Processing III*; Society of Photo Optical: Bellingham, WA, USA, 1988; pp. 186–208.
70. Cheng, Y.W.; Sun, P.C.; Chen, N.S. The essential applications of educational robot: Requirement analysis from the perspectives of experts, researchers and instructors. *Comput. Educ.* **2018**, *126*, 399–416. [CrossRef]
71. Cheung, A.C.; Slavin, R.E. The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educ. Res. Rev.* **2013**, *9*, 88–113. [CrossRef]
72. Slavin, R.; Madden, N.A. Measures inherent to treatments in program effectiveness reviews. *J. Res. Educ. Eff.* **2011**, *4*, 370–380. [CrossRef]
73. Strelan, P.; Osborn, A.; Palmer, E. The flipped classroom: A meta-analysis of effects on student performance across disciplines and education levels. *Educ. Res. Rev.* **2020**, *30*, 100314. [CrossRef]
74. Belpaeme, T.; Baxter, P.; De Greeff, J.; Kennedy, J.; Read, R.; Neerincx, M.; Baroni, I.; Looije, R.; Zelati, M.C. Child-robot interaction: Perspectives and challenges. In Proceedings of the 5th International Conference on Social Robotics, Bristol, UK, 27–29 October 2013; pp. 452–459.

75. Belpaeme, T.; Kennedy, J.; Ramachandran, A.; Scassellati, B.; Tanaka, F. Social robots for education: A review. *Sci. Robot.* **2018**, *3*, eaat5954. [CrossRef]
76. Fernández-Llamas, C.; Conde, M.A.; Rodríguez-Lera, F.J.; Rodríguez-Sedano, F.J.; García, F. May I teach you? Students' behavior when lectured by robotic vs. human teachers. *Comput. Hum. Behav.* **2018**, *80*, 460–469. [CrossRef]
77. Hashimoto, T.; Kato, N.; Kobayashi, H. Development of educational system with the android robot SAYA and evaluation. *Int. J. Adv. Robot. Syst.* **2011**, *8*, 28. [CrossRef]
78. Serholt, S. Breakdowns in children's interactions with a robotic tutor: A longitudinal study. *Comput. Hum. Behav.* **2018**, *81*, 250–264. [CrossRef]
79. Whittier, L.E.; Robinson, M. Teaching evolution to non-English proficient students by using Lego robotics. *Am. Second. Educ.* **2007**, *35*, 19–28.
80. van Straten, C.L.; Peter, J.; Kühne, R. Child-robot relationship formation: A narrative review of empirical research. *Int. J. Soc. Robot.* **2020**, *12*, 325–344. [CrossRef] [PubMed]

education sciences

Article

Plant Identification in the 21st Century—What Possibilities Do Modern Identification Keys Offer for Biology Lessons?

Alexander Finger [1,*], Jorge Groß [2] and Jörg Zabel [1]

[1] Institute for Biology, Biology Education, Leipzig University, 04109 Leipzig, Germany
[2] Didactics of Natural Sciences, Phillips University of Marburg, 35037 Marburg, Germany
* Correspondence: alexander.finger@uni-leipzig.de

Abstract: The loss of biodiversity and the accompanying "plant blindness" are major problems for mankind. Biology classes in Germany deal with this topic with the aim of enabling students to identify plants in their surroundings. Here, the process of plant identification plays a key role. To render the process of plant identification, more student-oriented, new digital approaches are being developed. Thus, teachers are now being confronted with digital tools for plant identification without having exact knowledge of their added value. This intervention study was therefore conducted in order to determine the effects on learning by means of a paper-based dichotomous identification key (Eikes Baumschule) and a digital identification app (ID-Logics). The results show that both tools have individual media-related differences that should be considered when designing learning strategies: With the previously reduced, paper-based tool, students can identify plants more quickly and often more correctly. However, the digital app has advantages in terms of enjoyment and learning about individual characteristics of plants. The study shows the challenges and opportunities associated with the (digital) medium. Furthermore, it sheds light on the process of species determination and reveals further fields of research in science education.

Keywords: media in education; mobile learning; teaching/learning strategies; species identification

Citation: Finger, A.; Groß, J.; Zabel, J. Plant Identification in the 21st Century—What Possibilities Do Modern Identification Keys Offer for Biology Lessons? *Educ. Sci.* **2022**, *12*, 849. https://doi.org/10.3390/educsci12120849

Academic Editors: Jian-Hong Ye, Yung-Wei Hao, Yu-Feng Wu and Savvas A. Chatzichristofis

Received: 24 October 2022
Accepted: 15 November 2022
Published: 23 November 2022

1. Introduction

The loss of biodiversity is one of the biggest challenges of the 21st century [1]. But it seems that this problem is not a priority in the public eye. This low awareness can partially be explained by the lack of species knowledge in society. Whereas large animal species are still in the focus of social perception, the decline in inconspicuous animal species is hardly registered. Empirical studies show that knowledge and interest about plants is anything other than widespread. This phenomenon is called "plant blindness" [2]. Studies have shown that this problem has been increasing in recent years [3,4]. Numerous studies have attempted to combat plant blindness by investigating and developing new methods for improving plant knowledge in schools, e.g., [5,6]. These studies have shown that young students have a poor or very limited knowledge of species, and that conveying or improving this knowledge is a difficult task for educators.

One recent hope for motivating students to take more interest in plants and their identification lies in the development and use of digital identification tools. Many scientists are advocating the digital approach towards the identification of plants, since they expect an easier accessibility as well as better learning outcomes [7–9]. But only few empirical studies have so far analysed the effects of digital identification tools on students. Stagg and Donkin [10] conducted a comparison study in the United Kingdom using the usability concept, in order to test the effectiveness of printed identification keys versus digital multi-access keys. The results of the study were inconclusive as to which key was more usable. But they also speculated that younger students may show a stronger preference for the digital application. Jacquemart [11] tested a digital dichotomous identification key in basic level university courses by creating a distance learning environment. This setting revealed that the digital presentation of plants can be useful for students to practice their identification skills and gather knowledge about plants and their characteristics.

But there is still little known about the impact of digital tools at school level and whether they can improve learning processes. In general, empirical information on the impact of digital media on student learning is limited. So far, digital tools have often been studied under lab conditions and results have been obtained that are difficult to transfer to authentic field trip contexts in schools. Therefore, teachers and professionals know too little about the impact of digital tools, particularly when they are used in authentic learning environments. Nevertheless, they are obliged to use them, since some German federal states such as Saxony-Anhalt included the identification of plants with digital tools in the biology curriculum of the 9th grade [12].

But what potential do modern plant identification keys designed for schools really offer, and how do they support the learning of plant identification and specific plant characteristics? This question remains open in view of the inconclusive results and the few studies available. To answer this question, we investigated the effects of two different identification tools: a digital polytomous key and a reduced paper-based dichotomous key for plant identification. In order to provide an authentic context for our study, data were collected during a school field trip.

2. Theory

2.1. Plant Blindness and Species Knowledge

Bollnow [13] (p. 119) stated: 'We see only those plants for which we have names'. Accordingly, an aspect of plant blindness is a lack of knowledge about plants (species knowledge) and another is the inability to recognise individual plant specimens by means of species identification [14]. In a broader context, the knowledge of species is thus necessary as a basis for understanding biodiversity. Only through species knowledge and the appreciation of biodiversity can problems such as biodiversity loss be recognised by students [15–18]. Hooykaas et al. [19] presented the concept of 'species literacy'—a concept of biodiversity which focuses on species—in order to tackle the loss of biodiversity. Therein, the ability to identify species and thus the understanding of the species concept are described as essential in order to increase people's awareness of biodiversity and species loss. Despite their great significance for life, plants have a very low priority for school children [20,21]. In Germany, for example, the interest of young people in identifying plants and in the subject area of botany halved from 40% to 20% between 1997 and 2003 [22]. International studies also indicate that plants play only a subordinate role in students' interest [23]. Subsequently, it is not surprising that students only have a very rudimentary knowledge of species [24]. Jäkel and Schaer [17] were able to show that primary school students can name less than five species with certainty. Especially in the case of woody plants, great deficits in species knowledge were revealed. Further studies demonstrate that students recognise only a few species in general, such as dandelions and daisies [15,17,24–28]. Tunnicliffe [29] was able to demonstrate that primary school children strive to learn or know the names of the plants presented to them in order to match what they see with their own concepts. The studies from Lindemann-Matthies [30] or Jäkel and Schaer [17] also showed that both interest and species knowledge can be positively influenced by intensive engagement with the subject in school and out-of-school learning situations. Especially out-of-school environments that cannot be replicated in the classroom seem to be an effective way of increasing student interest and engagement in a subject [31]. As has already been shown in other studies on interest in plants, a subject-specific difference between girls and boys exists in this respect [27,32,33]. This gender difference also influences the ability to recognise species. For example, various studies have shown that female students have a better knowledge of plants than male students [30,34,35]. On the one hand, this manifests itself in the general knowledge of species, and on the other, in the ability to recognise them. Suggested possible influencing factors for this difference are a higher interest in animals and plants [15,17] or in biology lessons in general [21], as well as a higher appreciation for plants [35]. It is this literature that roused our interest in the gender-specific effect of the plant identification tools on species learning.

2.2. Identification Tools

Educational researchers have long been looking for opportunities to address the problem of increasing plant blindness in students and to find the corresponding answers. Contact with nature and plants should be encouraged in an early stage of child development, such as in kindergarten and elementary school, e.g., [26]. Interest in plants, however, often diminishes as personal development progresses, as plants and nature do not have a prominent standing in our technically oriented world [23]. One reason for the growing plant blindness among students seems to be that nowadays, students rarely experience nature directly. Instead, digital media usually play a dominant role in students' everyday life. According to the Bitkom study [36], 67% of 10- to 11-year-olds and already 92% of 14- to 15-year-old students own a smartphone. Consequently, it makes sense to address young people through this medium. As a result, many applications and approaches have been developed in recent years for plant identification. Some of these applications use pattern recognition to virtually automatically identify plants based on digital image recognition. Chaki et al. [37] developed a Neuro-fuzzy classifier that uses plant leaves, which proved an 80–100% classification accuracy, depending on the respective species. Sekeroglu and Inan [38] demonstrated that their neural network approach could recognise 97.2% of the 27 leaf types. Often, these programs use images of distinct plant features such as leaves or blossoms to identify a species. All of these popular programmes take pictures and transmit them online to a server for analysis, which performs the identification, so that in the end the species can be identified by the program.

But botanical species expertise in school is not only focused on memorising the correct names of species. Rather, a (blind) dependence on a purely server-based identification should be counteracted. And a competence for a scientifically oriented education should be built up among the students. The main objective here is to train students in one of the oldest skills in biology: the ability for plant identification. For learning and identification in the case of living beings, it is crucial to take students by the hand and train them to identify essential biological characteristics. Through the ability to identify plants, students are encouraged to take a closer look at these, and discover the unique biological characteristics of individual species, which could help reduce the students' plant blindness. Also, enhancing species knowledge is assumed to increase the awareness for biodiversity and its possible loss [19]. Therefore, these identification tools that seem promising and appealing at first sight, and work with an automatic identification algorithm were not considered in this study, since the learning goals do not focus only on the species name. Thus, this study focused on two identification keys specifically designed for school use.

Dichotomous paper-based identification key

Conventional identification tools such as the identification key 'Rothmaler–Excursion Flora of Germany' [39] are built up dichotomously. This means that users must decide between two descriptions of a certain identification character of the plant to move forward in the process. The identification process is predetermined and does not consider season or location. In order to use these tools, students require extensive knowledge about scientific botanical terms, which are used to describe a certain characteristic of a plant and its different appearances. Furthermore, certain characteristics are seasonal and may not be available at the time of the identification process. This may lead to problems and result in the failure of the process, which is accompanied by demotivation and loss of interest. Such identification tools are mostly used by experts and are rarely applicable in schools. To give students an easy access and the ability to determine plants in a specific area, different tools have been developed in recent years. For instance, Feketitsch [40] developed an adaptable dichotomous identification key based on a digital algorithm. This digital identification tool allows teachers and educators to create a printable key before beginning the lesson with a reduced scope of species—specific for the shrubs and trees of a certain area. In order to create a customised printable version for plant identification, an educator can select a subset of species out of a total of 85 shrubs and trees in the tools list from Eikes Baumschule [41]. The identification process is guided and based on the graphical representation of the important characteristics (Figure 1). It reduces the botanical terms and written descriptions of the plant features to a minimum, so that laypersons can easily identify shrubs and trees within a limited spectrum of species. Nevertheless, these advantages (easy access, the limitation of species scope) mean that the tool cannot be used universally in every situation, since a preselection of the species is required.

Polytomous multi-access key

A rather new possibility for identifying plants that opens up from digitisation is the use of polytomous identification keys. The ID-Logics app (INITREE Software GmbH, Berlin, Germany) used in this study was designed and developed specifically for students under consideration of their problems in understanding written descriptions. The digitisation provided developers with new ways of presenting plants and guiding students through the identification process. Here, the dichotomous path was replaced by a polytomous algorithm which presents more than two characteristics of a plant feature to the user in the form of a graphic representation (Figure 1). For previously identified problems with the recognition of plant characteristics, explanatory videos are offered that establish a connection between form and function [42]. In addition, the app possesses a fault tolerance that helps students in the process of identification. In this way, the algorithm of the ID-Logics app can determine the probability of remaining plants and quickly reduce the number of species. Only specific characteristics are presented that usually lead to about 1–6 remaining final species. These species are presented as pictured and illustrated fact sheets to help the students verify the plant in front of them. By using digital media with access to the current date, the season is also taken into consideration for the identification process. Thus, the ID-Logics app only offers features that are available during the particular season.

Similarities and differences between the identification tools

Nowadays, teachers are faced with the question of which tools are suitable for identifying plants, as they can choose from digital and analogue tools. The use of digital tools, for example in plant identification, opens up new possibilities which, according to the SAMR model [43], should go beyond a simple digitisation of an analogue structure, i.e., a simple replacement, in order to offer an added value process (Table 1). For example, a digitally designed tool such as ID-Logics not only enables new approaches towards plant identification, but also presents teachers with new opportunities and challenges in its use. This leads to a redefinition of the learning process [43]. Therefore, the intention of this paper is to investigate the impact of a new digital approach with a similar identification tool. As already mentioned, the digital assessment process cannot be reproduced on paper, so that this attempt also has its limitations. Nevertheless, an attempt was made to ensure a high degree of comparability in the choice of tools. Thus, both keys were developed empirically for the use by students to help identify plants [40,42]. Both tools focus on a graphical representation during the identification process rather than on text-based descriptions. During the development process of both tools, the developers (Affeld & Groß: ID-Logics; Feketitsch & Lehnert: Eikes Baumschule) reported that novices indeed benefit from this approach. But in some aspects, the tools differ: While Eikes Baumschule is a dichotomous, paper-based identification tool that offers a pre-selected species spectrum, the ID-Logics app is a polytomous multi-access key that works with the full range of species (n = 192) on offline digital devices.

Table 1. The SAMR model for assessing the integration of digital technologies in biology education adapted from [43,44].

	Level	Definition	Added Value
Enhancement	Substitution	An analogue medium is digitally offered.	No functional enhancement
	Augmentation	Digital technologies provide improvements (e.g., audio, video and animation).	Low functional extension (e.g., non-linear information access).
Transformation	Modification	Digital technologies bring about fundamental changes in teaching (e.g., cooperative location-independent work with etherpads).	Individualised, cooperative and product-oriented-constructive extension.
	Redefinition	Use of digital technologies enables new forms of teaching (e.g., non-dichotomous identification of plants).	Teaching and learning opportunities that cannot be implemented with analogue media.

2.2.1. Usability of an Identification Key

Identification keys are important tools for laypersons and experts in order to identify species and learn about nature [45]. These keys are also important for estimating biodiversity and, moreover, its loss. In 1992, the Global Biodiversity Assessment stated that 'The range of available field guides, keys and other identifying aids is a major constraint to the assessment of biodiversity' [46] (p. 568). To study biodiversity with students despite this constraint, identification keys with high usability should be used. This may motivate students to use these tools in their everyday life as well as enhance their knowledge about nature. The usability concept was defined by Lawrence and Norrish [47] with three parameters: efficiency, satisfaction and effectiveness. Efficiency is the minimised time and effort required to identify a plant. Satisfaction is the enjoyment of using the guide. Effectiveness is the ability of a user to make a successful identification. Since success is not the only parameter for effectiveness in school, the concept was expanded in this study to include species knowledge and characteristics learning.

2.2.2. Characteristics Learning—The 'Leaf or Leaflet' Problem

The study from Affeldt, Groß and Stahl [48] showed that students can generally identify several tree characters, but tend to focus on the description of leaves. The 'leaf or leaflet' case chosen as an example illustrates the following educational challenges: Most of the students' conceptions of leaves differ from the scientific perspective. In the study from Affeldt, Groß and Stahl [48] (teaching experiments, $n = 28$), students described all kinds of leaves as 'green', 'simple in shape' and 'with a stalk'. Botanists also use leaves as a central character to identify trees. In doing so, they distinguish between simple, pinnate or lobed leaves. The ability to differentiate between different types of leaves is a prerequisite for conventional identification tools, but most students have problems in distinguishing between simple leaves and the leaflets of pinnate leaves. The study demonstrated that students were not aware of these differences and interpreted a leaflet as a simple leaf. The authors pointed out that these students used their basic-level concepts of leaves [48]. These concepts reflect our everyday experiences and disregard complex issues of scientific conceptions and therefore, most of the students do not apply the concept of pinnate leaves. As a consequence, in the process of tree identification, students encounter situations in which they need assistance. In the ID-Logics-app, two ways of meeting these challenges were designed: Firstly, interventions with additional information were developed. In case of the 'leaf-or-leaflet' example, an explanatory video was used that explains how to identify a single leaf: 'You can identify a single leaf by looking at the connection of the base of the leaf-stalk to the twig: it broadens and there is a bud directly above it.' (ID-Logics). Secondly, the database was enlarged with regard to the students' perspectives. Trees with pinnate leaves now contain the trait 'simple' in the feature library in the database. Based on these changes, the process of tree identification is more likely to end with the scientifically correct naming of the species. In contrast to traditional identification tools, the diversity of features and the quantity of identification steps during the process were reduced. Hence, a programmatic logic with multifactorial reduction in characters was designed, which uses information about the remaining species and available features. Eikes Baumschule also aims at taking the pupils' concepts into account in the identification key. Similar to the ID-Logics app, it uses mainly symbolic representations of plant features. However, to overcome misconceptions in the identification process, it rather relies on the teacher. In the identification key, only graphic accents such as a red outline of features are offered.

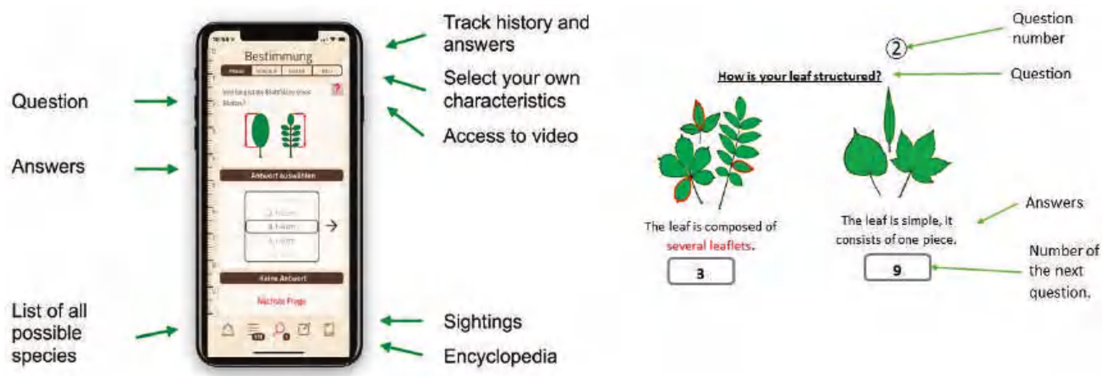


Figure 1. The app ID-Logics (**left**) [49], dichotomous identification key Eikes Baumschule (**right**) [41].

2.3. Research Question and Hypotheses

The central question of the exploratory study was whether modern identification media represent a more effective and efficient approach to the well-known problem of plant blindness. In other words, to what extent does a digital plant identification key help students learn plant identification and specific plant features? In order to empirically investigate this rather broad question, the concept of usability [47] was adopted and the question was divided into several hypotheses.

Hypothesis 1. *Effectiveness of the tool.*

We expect both the polytomous multi-access key (ID-Logics app) and the reduced identification key (paper based) to positively influence the ability to identify plants. Furthermore, the type of identification key also has an influence on the recognition of the leaflet. Female students are expected to remember more species correctly than male students.

H1a. *Due to the visual scaffolding we expect that the ID-Logics app should be more effective than the reduced identification key.*

H1b. *Due to the graphical presentation and the assistance in the ID-Logics app designed to meet learner's needs, we assume that students will learn to recognise morphological features such as the leaflet better with the ID-Logics app.*

H1c. *Female students can remember more species correctly when using the ID-Logics app, due to a higher affinity for plants.*

Hypothesis 2. *Efficiency of the tool.*

The type of plant identification key has an influence on the time required.

H2. *Due to its limited number of species, we assume that the reduced identification key will lead to a quicker identification process.*

Hypothesis 3. *Satisfaction with the tool.*

The type of plant identification key influences the reported enjoyment.

H3. *Due to the graphical presentation, the individual support functions and the general enthusiasm of the students for technology, we assume that the ID-Logics app will achieve a higher level of enjoyment than the reduced identification key.*

3. Materials and Methods

3.1. Participants and Procedure

This experimental study was carried out in collaboration with a high school in Saxony-Anhalt, Germany. The intervention consisted of a one-day field-trip to a meadow with typical shrubs and trees. All students were in the late 9th or early 10th grade and were on average 14–16 years old (M = 14.8, SD = 0.628). The intervention was planned for 310 students from a total of twelve classes. Of these, 283 (52% male, 48% female) took part in the study, spread over four field-trip days.

The respective sample size of the study varies depending on the research design and question and is therefore shown separately in the presentation of the results. Missing values that arose during the research are analysed in the Section 6.

The students were informed about the field-trip 14 days beforehand by the first author and their biology teachers. Prior to the field-trip, the students had received no specific training or preparation for identifying plants. The students went through the normal school curriculum, which focused on plant and species identification in the 7th grade. Since the field-trip was part of the regular biology class, all students were required to participate.

During the field-trip, the students visited eight learning stations that focused on different aspects of the meadow's ecosystem (e.g., water quality, calculation of the flow velocity of the river Saale, soil analysis). Since the field-trip area was located on a small island, the students assembled into groups of 10 people at maximum and walked freely between the stations. To prevent agglomerations at the stations, the groups were advised to move clockwise between the stations. Thus, each station was only visited by one group at a time and no direct interaction between the groups was possible. One of the learning stations was the plant identification station. Here, each student had to identify five previously marked shrubs and tree species. Two of these species possessed compound leaves comprising of leaflets, which the students had to consider during the identification. To evaluate the usability of the assigned tools during the plant identification exercise, an experimental setting was employed. The student groups were randomly assigned to one of two tools (see Figure 2). They either received an iPad with the identification software ID-Logics or a paper-based identification tool derived from Eikes Baumschule (see Section 2.2). The paper tool was designed for the specific area and only included 25 local tree and shrubs species. After assigning the groups to the tools, standardised instructions were given, specific to the tool used. Arriving at a tree, students were asked to write down their assumption of the tree's species name. Thereafter, they commenced the identification process. Each time the students identified a plant, they were asked to note their identification results. The students could decide on the order of the identification process themselves and were allowed to move freely within the plant identification station area. No time limit for the identification process was given. The time the students needed to identify a tree species was recorded by a research assistant who helped with general orientation in the area but was not allowed to intervene during the process. In order to measure their satisfaction, the students were to fill out a questionnaire immediately before and after they entered/left the identification area. The standardised questionnaire focused on intrinsic motivation and contained an enjoyment scale.

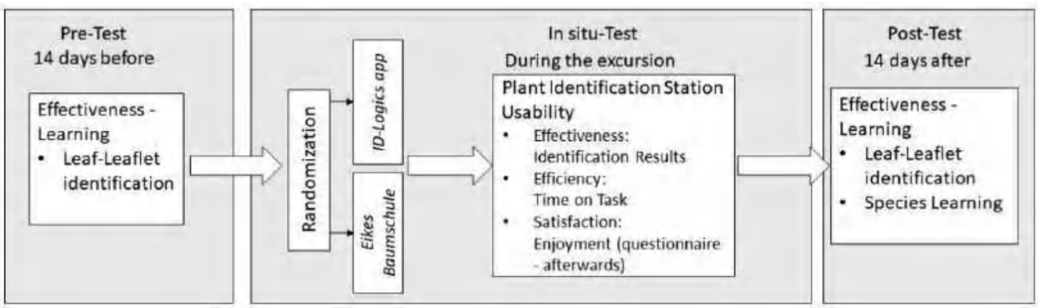


Figure 2. Study design.

3.2. Data Collection

To answer the research question, we determined the following variables: usability, characteristic identification, and species learning. Data were collected before, during and after the field-trip by means of observation as well as via pen-and-paper questionnaires (see Figure 2).

3.2.1. Usability

Students were asked to note their final result after each identification process. To evaluate the usability of the two keys, the effectiveness, efficiency and satisfaction were measured according to Lawrence and Norrish [47]. In order to assess the effectiveness of the identification tools, the students' success rate was determined as the proportion of correctly identified species with scores ranging from 0 to 1.

To measure the efficiency of the identification tools, the time needed for the identification process was recorded. The students' satisfaction with the key was measured using a short questionnaire and the 'enjoyment' scale, provided after the identification procedure. The 'enjoyment' scale was taken from the Intrinsic Motivation Short Scale (SIM), which is based on the theories and templates on intrinsic motivation by Decy and Ryan [50]. The SIM was chosen because it is designed for school situations and has already been translated into German [51]. Studies also showed that the scale 'enjoyment' can be deemed reliable and valid (Table 2). The students were instructed to rate their agreement to the respective statements on a 5-point-Likert-scale (ranging from 1 = I do not agree to 5 = I agree completely). For further information see the results that will be published separately in Finger, Bergmann-Gering and Groß [52].

Table 2. Background information on the used self-report scale 'enjoyment'.

Scale	Items	Example	Origin	Cronbach's α	Based on	Used by
enjoyment	3	I enjoyed identifying species using the identification tool	[50]	0.77	[50,53–55]	[56–58]

3.2.2. Characteristic Identification Test—Leaf or Leaflet

In order to assess if the tools can support the learning of certain plant characteristics, the students' ability to learn and recognise an important characteristic was measured. For this purpose, students were examined 14 days before (pre-test) and 14 days after the field-trip (post-test) (Figure 2). During the test, they were asked to circle a single leaf of each species presented. Of the 4 presented species, only one had the leaflet characteristic, the others had leaves.

3.2.3. Species Learning

To test the effect of the tools on species learning, we evaluated the students' ability to remember the identified plants during the post-test 14 days after the field-trip (Figure 2).

The students were required to write down the five shrubs and trees which had been presented to them during the field-trip. The number of correctly named species (0 to 5) was then recorded for each student.

3.3. Data Analysis

All data analyses were performed with IBM SPSS Statistics for Windows 25.0 (IBM Corp., Armonk, NY, USA). We used different analytical approaches with regard to the respective hypothesis. Assumptions (i.e., no significant outliers, normal distribution of residuals) were checked in the case of all analyses by means of extensive descriptive analyses. Unless explicitly stated in the Section 4, all assumptions were met.

Group-specific differences in effectiveness (rate of successful plant identification), efficiency (time taken) and enjoyment (self-report scale) were each analysed with a One-way Analysis of Variance (ANOVA) with the medium (digital or paper-based) as a fixed effect (H1a, H2, H3).

In order to examine the relation between the medium used and the ability to identify a leaflet before and after the intervention, a repeated-measurement ANOVA was conducted, with time (pre/post intervention) as within-subject factor, medium as between- subject factor, and expected interactions between both factors (H1b).

Differences in the students' ability to learn species names (H1c) were analysed using a two-way ANOVA with gender (male/female) and medium (digital/paper-based), as well as their interaction as fixed effects. The pairwise comparisons of the subgroups for medium and gender were performed using a post hoc test with Bonferroni correction of *p*-values.

4. Results

4.1. Usability

The results on effectiveness revealed a large and statistically significant difference between the tools ($F(1,217) = 56.32$, $p < 0.001$, partial $\eta^2 = 0.206$, Table 3). The efficiency of the tools was determined by recording the individual identification time. It differed statistically significantly with a large effect between the two tools ($F(1,159) = 23.53$, $p < 0.001$, partial $\eta^2 = 0.275$). Students who used the digital tool were engaged in the process for about 6 min longer than the group with the reduced paper-based identification key. The results on enjoyment showed a statistically significant, but small-sized effect between the tools ($F(1,250) = 6.748$, $p < 0.01$, partial $\eta^2 = 0.026$.

Table 3. ANOVA results for average success and efficiency of the plant identification.

Key Characteristic	Variable	Reduced Paper-Based Identification Key		Digital Identification Key (ID-Logics App)		F	*p*	N^2	Partial η^2
		M^1	SD	M^1	SD				
effectiveness	Proportion of correctly identified species	0.76	0.27	0.49	0.27	56.32	<0.001	218	0.206
efficiency	Total time of species identification [min]	18.98	1.166	25.139	0.885	17.704	<0.001	160	0.275
enjoyment	Enjoyment	3.15	0.088	3.48	0.091	6.748	0.01	251	0.026

M^1 (mean), SD (standard deviation). N^2 (sample size), partial η^2 (effect size).

4.2. Characteristic Identification Test—Leaf or Leaflet

Table 4 shows the test results in relation to the learning outcomes when differentiating between leaf and leaflet. To avoid falsification from a pre-test effect, the survey was carried out in a pre- and post-test inquiry on different tree species and was statistically evaluated [59].

Table 4. Descriptive statistic for leaf/leaflet recognition (% correct and false answers) (N = 197).

	Pre-Test							
	Reduced paper-based identification key				Digital identification key (ID-Logics app)			
	European Beech	Sycamore Maple	Common Hazel	Robinia	European Beech	Sycamore Maple	Common Hazel	Robinia
Leaf/leaflet	Leaf	Leaf	Leaf	Leaflet	Leaf	Leaf	Leaf	Leaflet
Correct	100	99.21	100.00	18.25	100.00	100.00	100.00	13.60
False	0.00	0.79	0.00	81.75	0.00	0.00	0.00	86.40
	Post-Test							
	Reduced paper-based identification key				Digital identification key (ID-Logics app)			
	Pedunculate Oak	Norway Maple	Hornbeam	Common Ash	Pedunculate Oak	Norway Maple	Hornbeam	Common Ash
Leaf/leaflet	Leaf	Leaf	Leaf	Leaflet	Leaf	Leaf	Leaf	Leaflet
Correct	97.62	99.21	97.33	30.16	98.40	96.80	97.60	60.00
False	2.38	0.79	2.67	69.84	1.60	0.80	2.40	40.00

This learning outcome was statistically significantly different with a medium effect size between the two tools used by the students only for the leaflet condition, $F(2,195) = 21.3$, $p < 0.000$, partial $\eta^2 = 0.098$). Figure 3 shows the proportion of correctly identified leaflets in the pre- and the post-test for both tools.

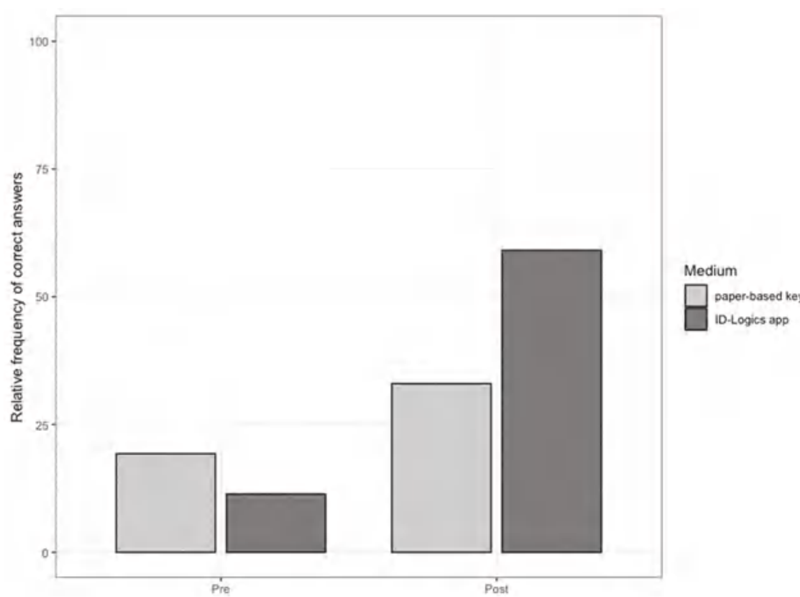


Figure 3. Leaflet recognition pre- and post-test (N = 197).

4.3. Species Learning

The type of identification tool used by the students had no significant effect on the students' ability to remember the identified species ($p = 0.88$). Students were able to remember 3.40 species on average when using the ID-Logics app and 2.88 species when using the paper-based identification tool ($F(1,235) = 2.943$, partial $\eta^2 = 0.020$).

In contrast, gender had a statistically significant influence ($p < 0.001$) on the students' ability to remember the identified species (Table 5). In general, female students performed better than male students, regardless of the tool used (mean female: 3.790; mean male: 2.486).

Table 5. ANOVA results for remembering the identified plant species.

	Reduced Paper-Based Identification Key				Digital Identification Key (ID-Logics App)				p_{female}	p_{male}	d_1[2]	d_2[3]
	Female		Male		Female		Male					
	M[1]	SD	M	SD	M	SD	M	SD				
Rem. Spec.[4]	3.76	1.79	2.00	1.79	3.82	1.79	2.97	1.79	0.889	0.028	−0.033	−0.541

M[1] (mean), SD (standard deviation). d_1[2] (Cohens d for comparison of tools in the female subgroup). d_2[3] (Cohens d for comparison of tools in the male subgroup). Rem. Spec.[4] number of correctly remembered species [0 to 5].

Overall, the interaction of medium and gender was not significant ($p = 0.13$). On average, females performed better than males and the results between the two tools used did not differ significantly (Mean females: paper-based tool = 3.76, digital tool = 3.82). However, male students performed significantly better when using the digital tool than with the paper-based tool ($F(1,235) = 4.90$, $p = 0.028$, partial $\eta^2 = 0.033$ (Mean: paper-based tool 2.00, digital tool 2.97)). Here, the tool used had a small effect size.

5. Discussion

The aim of this explorative study was to investigate the effects of two different identification tools (digital polytomous and reduced paper-based dichotomous) using the adapted usability concept under realistic field conditions. The usability concept was expanded in the effectiveness category to include the two indicators of determining characteristics (leaf or leaflet) and learning species names. These additions were made in order to better assess the applicability for biology teaching.

Our investigation revealed interesting results: In terms of effectiveness, data show that students who worked with the reduced identification key were able to correctly identify more species on average. In terms of learning, the study shows that students using ID-Logics were more likely to correctly identify the leaflet characteristic. The data reveal that the tool used had no significant influence on species learning. However, in a subgroup analysis we found that male participants remembered more species when they used ID-Logics, while the tool made no difference in the case of females.

In terms of efficiency, the results show that students using the reduced identification key were able to identify plants faster. As to enjoyment, students using ID-Logics reported higher ratings.

Overall, both tools in our study have the potential to support students with the difficult process of plant identification in the field. However, our findings suggest different opportunities and challenges for each of the two tools. We will discuss the educational implications of these findings in the Section 7.

5.1. *Effectiveness of the Tool*

The first hypotheses concerned the effectiveness of the tool. H1a: 'Due to the visual scaffolding the ID-Logics app should be more effective than the reduced identification key'. Indeed, our results indicate that there is a significant difference between the two tools regarding the effectiveness. Students who used the ID-Logics app succeeded in correctly identifying a plant in 56% of the attempts, while the group using the paper-based key had a 76% success rate. Therefore, we fail to reject the null hypothesis 1a. Moreover, hypothesis 1a has to be revised to: 'Students who have used the reduced paper-based identification key can correctly identify more plants than students who have used the polytomous digital identification tool'. One explanation for this result may be the smaller species spectrum of the reduced paper-based dichotomous identification tool (25 species). The ID-Logics app, on the other hand, is a more comprehensive identification tool with a total of 192 species. This may increase the possibility of errors and lead to a lower success rate when using the ID-Logics app, as more features had to be identified during the identification process. However, the final identification step at species level could also explain the observed difference, because with ID-Logics, students need to choose between several (up to six) very similar species. Hence, the risk of confusion is higher in this step compared to a dichotomous key. Also, the reduced identification key not only limits the total number of possible species, but also eliminates in advance similar species that do not occur locally but could be confused with those existent on site. This makes it easier for students to distinguish between species at the characteristic level, as there are fewer similar characteristics of closely related species to choose from. However, the comparison of similar species in the app can encourage a closer look at selected characteristics. This could also have an impact on the learning of characteristics, as in total, fewer differentiated characteristics are presented in the reduced determination key. As an indicator for this, the ability to distinguish between leaf and leaflet was measured.

H1b: 'Due to the graphical presentation and the learner-designed assistance in the ID-Logics app, we assume that students will learn to recognise the leaflet better with this tool'. The ability to recognise specific characteristics of plants is of great importance for the correct identification of species. In the case of shrubs and trees, the leaf in particular represents an important feature for identification. In the past, students had problems distinguishing between leaf and leaflet [42,48]. In this study, the leaflet was correctly identified by only 11% and 19%, respectively, of the students before the intervention (Figure 3). Students tend to assume that leaves are undivided [28,40]. The results of our study support the assumption of a misconception of leaves: in the pre-test, nearly 100% of the students were able to correctly identify the leaf but less than 20% could identify a leaflet (Table 4). Therefore, our results coincide with the study of Affeldt, Groß and Stahl [48] and indicate that students face problems when plants do not match their inexpert concept of an undivided leaf. After the intervention, we observed a significant difference between students who used the ID-Logics app and those who used the reduced paper-based key. A total of 59% of the students who used the ID-Logics app were able to identify the leaflet after the intervention, whereas only 28% of the students who used the reduced identification key could identify the leaflet correctly. Therefore, we fail to reject the null hypothesis 1b.

The question arises as to why the two tools had such a different effect. One possible explanation could be the built-in support systems of the digital tool: the ID-Logics app can help students overcome learning barriers by presenting explanatory videos and specific supporting information. Furthermore, the identification process itself possibly helped the students to distinguish the characteristic of the leaflet. Since the digital tool ID-Logics is based on multiple graphical representations of one characteristic, students can more easily compare the original plant with this representation in the tool [42]. These authors concluded that the feature of emphasising plant characteristics in the ID-Logics app enabled the students to learn these characteristics more efficiently.

Does it also help students to learn species names when they are familiarised with specific characteristics? Overall, our results show that the tool had no significant effect on the ability to remember the species identified during the intervention. However, students using the ID-Logics app tended ($p = 0.88$) to remember more species (3.40) than students with the reduced dichotomous paper-based identification tool (2.88). This tendency should be further investigated. In this context, hypothesis H1c was created: 'Female students can remember more species correctly when using the ID-Logics app, due to a higher affinity for plants'. The findings indicate that the tool only had a significant effect on species learning in the male student group (Table 5). This influence may be due to the generally higher affinity of male students for digital media [60]. Since interest is a relevant factor for learning, digital media in general, and the ID-Logics app in particular, offer an opportunity to encourage male students. Nevertheless, we fail to reject the null hypothesis H1c, since females could in total remember more species.

5.2. *Efficiency of the Tool*

To investigate the efficiency of the two tools, we formulated a second hypothesis (H2): 'Due to a limited species selection of the reduced identification key, we assume that the task of identifying five species can be solved more quickly.' We fail to reject the null hypothesis 2 (Table 3). Students using the ID-Logics app required on average six minutes longer for the identification process than those using the reduced paper-based key. One reason could also be the higher complexity of the ID-Logics app and thus, difficulties in using this digital tool. This could also lead to a more intensive preoccupation with the graphical representation and the identification process itself. The program ID-Logics has a non-linear structure and distinguishing characteristics are presented based on previous input. This may have required the students to focus on each and every step of the identification process, especially when starting the identification process from the beginning. This feature of the program could also have reduced the students' cognitive load (based on their pre-knowledge), since they were rather guided by the program, and this could even have resulted in a higher attention effect for the students. It is also known that increased attention has a positive effect on learning [61].

5.3. Satisfaction with the Tool

But does a general difference in the enjoyment level exist? To assess this, the last hypothesis (H3) was created: 'Due to the graphical presentation, the individual support functions and the general enthusiasm of the students for technology, we assume that the ID-Logics app achieves a higher level of enjoyment than the reduced identification key'. We fail to reject the null hypothesis 3. Results show a small but significant difference in the enjoyment level of the students after they finished the identification process. This could be explained by a general higher affinity for digital tools used in the learning processes or with the novelty effect. Since students were not familiar with any of the identification tools, this effect should theoretically be equal for both tools. But since the ID-Logics app was graphically more advanced and offered different support functions (such as explanatory videos), this may explain the higher attraction for the students. For a more detailed breakdown of the motivational as well as the enjoyment aspects of the intervention, see [52].

6. Limitations

Comparing applications in a realistic school environment, even more so in the context of a field trip, is difficult, which is why limitations can exist in the transferability of results. For example, not all environmental factors and variables that are held constant in the laboratory can be controlled in the field. Although both applications differ in the species range, they are methodologically based on similar principles; namely, the pictorial comparison of characteristics. Another approach towards plant identification is the textual representation of the characteristics, as used by many expert keys. Expert keys often use technical terms of such a specific nature that they can be a decisive obstacle in the identification process for students. They often need help, because they cannot easily overcome these hurdles [42,48,62].

Since this was an intervention that took place during school lessons and included a total of three independent measurement points spread over several weeks, it is obvious that not all students could be present at all measurement points, for example due to sick leave. Indeed, 27 of the 310 students in the sample group reported sick for the field-trip. In addition, 86 of the 283 students who attended the field-trip were reported sick at the pre- or post-test. We assume that this proportion of missing values was a completely random figure. Furthermore, from the 283 students who attended the field-trip, we only generated 218 datasets for our analysis of the correct identification of plants. This is because 65 students failed to submit their results at the end of the station. A relationship between the missing data and their expected answers to the questions cannot be assumed. Due to technical problems, in the end, the data from 160 students were finally available for evaluation for the efficacy analysis. Since this was a random problem, we do not expect that the missing values will have a significant influence on the validity of the results presented.

In general, this intervention study should help to investigate the use of identification tools in a realistic school setting. The study shows that further studies are required to evaluate the effects of identification tools on learning, as well as their possible use in the classroom. Due to this practical approach, limitations exist in terms of the instruments, indicators and design used, as described above.

7. Conclusions

This intervention study reveals how important it is to examine digital media in detail, with respect to their individual advantages and disadvantages for teaching. Our results indicate that, taking different aspects of usability into account, there is no simple truth or automatic preference for the one solution or the other. The results we obtained were unexpected, leading to the rejection of many of our hypotheses. But what can we conclude from these findings, particularly for educational practice and for further research?

7.1. Comparing Apples and Oranges?

First of all, it is relevant to note that our study design did not simply compare an analogue with a digital tool. Rather, with both tools, the first step is a digital one. In order to use Eikes Baumschule, the teacher first needs to specify all relevant tree species of the field-trip area, so that the program can generate a dichotomous key from this information. Therefore, teachers need to familiarise themselves with the local tree species as an initial step and thereby already perform part of the identification process themselves. Our results indicate that, based on this pre-selection, the students have certain correlative advantages in terms of effectiveness and efficiency compared to the polytomous multi-access key ID-Logics. Should not the faster and more effective tool be generally recommended for use?

The collected data for the self-assessment of intrinsic motivation (enjoyment) in accordance with the intrinsic motivation concept [50] show that here, the enjoyment of determining a plant does not necessarily correlate with the speed with which a task can be solved, since the perceived enjoyment was significantly higher when using the ID-Logics app. Other factors such as the aesthetics of the plant or the observation of plants have been identified as important influencing factors for interest/motivation development in other studies [63]. But to this purpose, time is required, because speeding through the tasks obviously has a negative effect. Also, the use of ID-logics can promote the learning of specific plant attributes relevant for identification, as it especially emphasises the aesthetics and the characteristics of plants. This has also been described in other studies as a positive factor in combating plant blindness [63].

Thus, we argue here that reduced, pre-selected dichotomous keys such as Eikes Baumschule are suitable especially for the first attempts in the identification process, as they enable quick successes and thereby reduce the feeling of uncertainty on the part of the users. This in turn should promote the sense of autonomy as well as self-determination [63,64]. In contrast, digital media such as the ID-Logics app can play out their advantages in a later, more advanced stage of the learning process. They offer digital support, explanatory videos and assistance at appropriate stages and thus help the learner individually to autonomously overcome difficulties in understanding. ID-Logics is particularly suitable for learning the features of plants such as the relationships between structure and function. Furthermore, according to our results, male students in particular could benefit from the use of ID-Logics, although they are often more indifferent towards plants [23]. This could be due to a generally higher affinity for digital tools, which generates higher attention and thus supports learning. Future research approaches should examine these interactions more closely.

7.2. Educational Implications

Considering the findings reported above, we conclude that a combined teaching strategy for plant identification is probably the best way to stimulate students' interest in nature and their awareness of plant blindness and biodiversity loss. The aim is to combine the advantages of both instruments: As a first step, the students should quickly experience success and gain confidence in the process of plant identification. For this purpose, a reduced conventional identification key presenting a manageable selection of species is a good choice. The students should be made aware at an early stage that this tool can only be used in certain areas and therefore has a limited range.

In a second phase, an advanced digital tool comparable with ID-Logics should be introduced, as it can be used universally and provides further information about the plants and their specific characteristics. This not only empowers the students to use an identification tool independently in their environment; with the additional information it provides, it also offers the possibility of expanding knowledge about the individual species and thus the understanding of biodiversity, which in turn could reduce plant blindness.

Another benefit of combining both identification tools is that it potentially facilitates the internal differentiation in the learning group. Students with learning difficulties could benefit from the limited offer of the reduced identification key, as they would not be overwhelmed by the selection of different features. High-achieving students can benefit from the openness of the ID-Logics app and its universal applicability. Furthermore, educators should not only focus on the quantity of species identified, but rather on the quality of the identification process. Students often have problems identifying too many plant species during a lesson. This leads to stress and confusion, so that the intensity of engagement with the individual plant, the intended learning effect and the students' motivation are diminished. Therefore, it is generally better to identify fewer species, but with a higher level of detail, and to build on these experiences productively [65].

7.3. Next Steps

The desired answer to the question of whether modern identification media are a new solution to help overcome plant blindness is still unknown. But our findings suggest that the question is perhaps oversimplified, and that we should focus not only on the media, but rather on what we wish to achieve with their use: it is not a general question about the device itself, but rather, which is the right tool for a specific learning goal in science education. Both types of tools present specific opportunities and challenges [66]. That is why further research is needed into how the use of digital media affects student motivation, interest and learning in different environments. Other tools not considered in this study, such as the automatic image recognition of plants, also represent a potential opportunity to motivate laypeople to identify plants with a low-threshold offer. Therefore, these tools should also be investigated in future research with regard to their effects on the usability for teaching.

This study reveals the great opportunities that digital media offer when it comes to plant identification. As the availability of digital platforms such as smartphones and/or tablet computers is no longer a limiting factor, digital tools provide a usable method of enabling students to identify plants. Nevertheless, more research is required to better exploit the potential of digital media and to constantly improve the digital approach. Only if educators know the limits and possibilities of these tools, will they be able to use them effectively and generate interest in plants, fight plant blindness and increase awareness of the global decline in biodiversity. For this purpose, teachers not only need further empirical data on the effectiveness of the applications, but also specific in-service training to make use of them.

Author Contributions: Conceptualization, A.F.; Methodology, A.F.; Formal analysis, A.F.; Investigation, A.F.; Writing—original draft, A.F.; Writing—review & editing, J.G. and J.Z.; Visualization, A.F.; Project administration, A.F. All authors have read and agreed to the published version of the manuscript.

Funding: The publication of this article was funded by the Open Access Publishing Fund of Leipzig University supported by the German Research Foundation within the program Open Access Publication Funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. IPBES. *Summary for Policymakers of the Global Assessment Report on Biodiversity and Ecosystem Services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*; Díaz, S., Settele, J., Brondizio, E.S., Ngo, H.T., Guèze, M., Agard, J., Arneth, A., Balvanera, P., Brauman, K.A., Butchart, S.H.M., et al., Eds.; IPBES Secretariat: Bonn, Germany, 2019.
2. Wandersee, J.H.; Schussler, E.E. Preventing Plant Blindness. In *The American Biology Teacher*; University of California Press: Berkeley, CA, USA, 1999; pp. 82–86.
3. Frisch, J.K.; Unwin, M.M.; Saunders, G.W. Name That Plant! Overcoming Plant Blindness and Developing a Sense of Place Using Science and Environmental Education. In *The Inclusion of Environmental Education in Science Teacher Education*; Bodzin, A.M., Shiner Klein, B., Weaver, S., Eds.; Springer: Dordrecht, The Netherlands, 2010; pp. 143–157.
4. Randler, C.; Osti, J.; Hummel, E. Decline in Interest in Biology among Elementary School Students During a Generation. *Eurasia J. Math. Sci. Technol. Educ.* **2012**, *8*, 201–205. [CrossRef]
5. Bromme, R.; Stahl, E.; Bartholomé, T.; Pieschl, S. The Case of Plant Identification in Biology: When is a Rose a Rose. In *Professional Learning: Gaps and Transitions on the Way from Novice to Expert*; Boshuizen, P.P.A., Bromme, R., Gruber, H., Eds.; Kluwer Academic Press: Dordrecht, The Netherlands, 2004; pp. 53–71.
6. Randler, C.; Bogner, F.X. Cognitive achievements in identification skills. *J. Biol. Educ.* **2006**, *40*, 161–165. [CrossRef]
7. Dallwitz, M.J.; Paine, T.A.; Zurcher, E.J. *"Interactive Identification Using the Internet." Towards a Global Biological Information Infrastructure–Challenges, Opportunities, Synergies, and the Role of Entomology*; European Environment Agency Technical Report; EEA: Copenhagen, Denmark, 2002; Volume 70, pp. 23–33.
8. Farr, D.F. On-line keys: More than just paper on the web. *Taxon* **2006**, *55*, 589–596. [CrossRef]
9. Drinkwater, R.E. Insights into the development of online plant identification keys based on literature review: An exemplar electronic key to Australian Drosera. *Biosci. Horiz.* **2009**, *2*, 90–96. [CrossRef]
10. Stagg, B.C.; Donkin, M.E. Apps for angiosperms: The usability of mobile computers and printed field guides for UK wild flower and winter tree identification. *J. Biol. Educ.* **2017**, *51*, 123–135. [CrossRef]
11. Jacquemart, A.-L.; Lhoir, P.; Binard, F.; Descamps, C. An Interactive Multimedia Dichotomous Key for Teaching Plant Identification. *J. Biol. Educ.* **2016**, *50*, 442–451. [CrossRef]
12. Kultusministerium Sachsen-Anhalt. Fachlehrplan Gymnasium–Biologie. 2016. Available online: https://lisa.sachsen-anhalt.de/fileadmin/Bibliothek/Politik_und_Verwaltung/MK/LISA/Unterricht/Lehrplaene/Gym/Anpassung_2022/FLP_Biologie_Gym_01082022_swd.pdf (accessed on 1 June 2022).
13. Bollnow, O.F. *Sprache und Erziehung*, 3rd ed.; Kohlhammer Verlag: Stuttgart, Germany, 1966.
14. Wandersee, J.H.; Schussler, E.E. Toward a theory of plant blindness. *Plant Sci. Bull.* **2001**, *47*, 2–9.
15. Goller, H. Kontextabhängiger Erwerb von Arten- und Formenkenntnissen im Biologieunterricht des Gymnasiums. Ph.D. Thesis, Universität Regensburg, Regensburg, Germany, 2002. Available online: http://www.bibliothek.uni-regensburg.de/opus/volltexte/2002/68/pdf/Endversion.pdf (accessed on 1 June 2022).
16. Hammann, M. Kriteriengeleitetes Vergleichen im Biologieunterricht. Ph.D. Thesis, Universität Kiel, Kiel, Germany, 2002.
17. Jäkel, L.; Schaer, A. Sind Namen nur Schall und Rauch? Wie sicher sind Pflanzenkenntnisse von Schülerinnen und Schülern? *Biol. Lehren Und Lern. Z. Für Didakt. Der Biol.* **2004**, *13*, 1–24.
18. Uno, G.E. Botanical literacy: What and how should students learn about plants? *Am. J. Bot.* **2009**, *96*, 1753–1759. [CrossRef] [PubMed]
19. Hooykaas, M.J.; Schilthuizen, M.; Aten, C.; Hemelaar, E.M.; Albers, C.J.; Smeets, I. Identification skills in biodiversity professionals and laypeople: A gap in species literacy. *Biol. Conserv.* **2019**, *238*, 108202. [CrossRef]
20. Löwe, B. Interessenverfall im Biologieunterricht. *Unterr. Biol.* **1987**, *124*, 62–65.
21. Löwe, B. *Biologieunterricht und Schülerinteresse an Biologie*; Schriftenreihe der Pädagogischen Hochschule Heidelberg, Bd. 9; Deutscher Studien-Verlag: Weinheim, Germany, 1992.
22. Brämer, R. Naturschutz contra Nachhaltigkeit? Jugendreport Natur 2003 zu den Folgen der Naturentfremdung. *Umwelt Bildung.* **2004**, *2*, 18–20.
23. Elster, D. Student interests–the German and Austrian ROSE survey. *J. Biol. Educ.* **2007**, *42*, 5–10. [CrossRef]
24. Hesse, M. Eine Neue Methode zur Überprüfung von Artenkenntnissen bei Schülern, Frühblüher: Benennen-Selbsteinschätzen-Wiedererkennen. In *Zeitschrift für Didaktik der Naturwissenschaften*; Springer: Berlin/Heidelberg, Germany, 2002; Volume 8, pp. 53–66.
25. Bebbington, A. The ability of A-level students to name plants. *J. Biol. Educ.* **2005**, *39*, 63–67. [CrossRef]
26. Jäkel, L. Alltagspflanzen im Fokus. Botanisches Lernen in Zusammenhängen–eine didaktische Herausforderung im Anfangsunterricht. In *Praxis der Naturwissenschaften–Biologie in der Schule*; AV Akademikerverlag: Saarbrücken, Germany, 2005; Volume 54, pp. 15–21.
27. Schussler, E.E.; Olzak, L.A. It's not easy being green: Student recall of plant and animal images. *J. Biol. Educ.* **2008**, *42*, 112–119. [CrossRef]
28. Lehnert, H.-J.; Benkowitz, D.; Feketitsch, D.; Jaun, B. Blätter, Blüten, Habitus: Wie pflanzliche Formenkenntnis erworben wird und was bei ihrer Vermittlung beachtet werden sollte. In *Theorie, Empirie, Praxis*; Mayer, J., Hammann, M., Wellnitz, N., Arnold, J., Werner, M., Eds.; Kassel University Press: Kassel, Germany, 2013; pp. 272–273.

29. Tunnicliffe, S.D. Talking about plants–comments of primary school groups looking at plant exhibits in a botanical garden. *J. Biol. Educ.* **2001**, *36*, 27–34. [CrossRef]
30. Lindemann-Matthies, P. Wahrnehmung biologischer Vielfalt im Siedlungsraum durch Schweizer Kinder. In *Lehr-und Lernforschung in der Biologiedidaktik*; Studien Verlag: Innsbruck, Austria, 2002; pp. 117–130.
31. Behrendt, M.; Franklin, T. A review of research on school field trips and their value in education. *Int. J. Environ. Sci. Educ.* **2014**, *9*, 235–245.
32. Klemm, I. Vergleichende Untersuchungen über den Pflanzenbegriff von Schülern. *Prax. Der Nat.* **1974**, *23*, 154–160.
33. Pohl, D.T. Naturerfahrung und Naturzugänge von Kindern. Ph.D. Thesis, Pädagogische Hochschule Ludwigsburg, Ludwigsburg, German, 2006. Available online: http://phbl-opus.phlb.de/frontdoor/index/index/docId/14 (accessed on 15 June 2022).
34. Bögeholz, S. *Qualitäten Primärer Naturerfahrung und ihr Zusammenhang mit Umweltwissen und Umwelthandeln*; Leske+Budrich: Leverkusen, German, 1999.
35. Lindemann-Matthies, P. Children's Perception of Biodiversity in Everyday Life and Their Preferences for Species. Ph.D. Thesis, Universität Zürich, Zürich, Switzerland, 1999.
36. BITKOM. Digitale Schule–Vernetztes Lernen. 2015. Available online: https://www.bitkom.org/sites/default/files/file/import/BITKOM-Studie-Digitale-Schule-2015.pdf (accessed on 11 October 2022).
37. Chaki, J.; Parekh, R.; Bhattacharya, S. Plant leaf classification using multiple descriptors: A hierarchical approach. *J. King Saud Univ. Comput. Inf. Sci.* **2018**, *32*, 1158–1172. [CrossRef]
38. Şekeroğlu, B.; Inan, Y. Leaves Recognition System Using a Neural Network. *Procedia Comput. Sci.* **2016**, *102*, 578–582. [CrossRef]
39. Jäger, E.J. Rothmaler-Exkursionsflora von Deutschland. In *Gefäßpflanzen: Grundband*; Springer: Berlin/Heidelberg, Germany, 2016.
40. Feketitsch, D. Eikes Baumschule–Entwicklung und Evaluation Einer Online-Bestimmungshilfe für Kinder. Ph.D. Thesis, Universität Karlsruhe, Karlsruhe, Germany, 2013.
41. Lehnert, H.-J. Eikes Baumschule. 2013. Available online: https://baum.bio-div.de/index.html (accessed on 15 June 2022).
42. Groß, J.; Affeldt, S.; Stahl, D. Find My Name! Evidence-based Development of an Interactive Species Identification Tool. In *Building Bridges Across Disciplines*; Eilks, I., Markic, S., Ralle, B., Eds.; Shaker: Aachen, Germany, 2018; pp. 97–108.
43. Puentedura, R.R. SAMR: Moving from Enhancement to Transformation [Web Log Post]. 2013. Available online: http://www.hippasus.com/rrpweblog/archives/000095.html (accessed on 20 September 2022).
44. Groß, J.; Lude, A.; Nerdel, C.; Paul, J.; Schaal, S.; Schmiemann, P.; Thyssen, C. Biologische Bildung in der Digitalen Welt. Die Digitale Transformation im Fokus der Biologiedidaktik. In *Fachliche Bildung in der Digitalen Welt*; Frederking, V., Romeike, R., Eds.; Waxmann Verlag: Münster, Germany, 2022.
45. Lawrence, A.; Hawthorne, W. Identifying Biodiversity: Why do We Need Field Guides? In *Plant Identification: Creating User-friendly Field Guides for Biodiversity Management*; Lawrence, A., Hawthorne, W., Eds.; Earthscan: London, UK, 2006.
46. Hawksworth, D.L. The Resource Base for Biodiversity Assessments. In *Global Biodiversity Assessment*; Heywood, V.H., Watson, R.T., Eds.; Cambridge University Press: Cambridge, UK, 1995; Volume 1140.
47. Lawrence, A.; Norrish, P. Producing a Successful Guide: Principles, Purpose, People and Process. In *Plant Identification*; Hawthorne, W., Lawrence, A., Eds.; Routledge: London, UK, 2006. [CrossRef]
48. Affeldt, S.; Groß, J.; Stahl, D. Die Artansprache verstehen–eine evidenzbasierte Analyse des Bestimmungsweges. *Lehr-Und Lernforschung Der Biol.* **2012**, *5*, 185–202.
49. Groß, J. ID-Logics, Bestimmung von Gehölzen. 2022. Available online: http://id-logics.com/ (accessed on 1 October 2022).
50. Deci, E.L.; Ryan, R.M. *Intrinsic Motivation and Self-Determination in Human Behavior*; Plenum: New York, NY, USA, 1985.
51. Wilde, M.; Bätz, K.; Kovaleva, A.; Urhahne, D. Überprüfung einer Kurzskala intrinsischer Motivation. *Z. Für Didakt. Der Nat.* **2009**, *15*, 31–45.
52. Finger, A.; Bergmann-Gering, A.; Groß, J. Does the medium matter?–The effect of a mobile digital identification tool on students' intrinsic motivation while identifying plants. *J. Biol. Educ.* **2022**, in press. [CrossRef]
53. Plant, R.W.; Ryan, R.M. Intrinsic motivation and the effects of self-consciousness, self-awareness, and ego-involvement: An investigation of internally controlling styles. *J. Pers.* **1985**, *53*, 435–449. [CrossRef]
54. McAurley, E.; Duncan, T.; Tammen, V.V. Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: A confirmatory factor analysis. *Res. Q. Exerc. Sport* **1989**, *60*, 48–58. [CrossRef] [PubMed]
55. Ryan, R.M.; Koestner, R.; Deci, E.L. Ego-involved persistence: When free-choice behavior is not intrinsically motivated. *Motiv. Emot.* **1991**, *15*, 185–205. [CrossRef]
56. Stavrova, O.; Urhahne, D. Modification of a school programme in the Deutsches Museum to enhance students' attitudes and understanding. *Int. J. Sci. Educ.* **2010**, *32*, 2291–2310. [CrossRef]
57. Wilde, M.; Hußmann, J.S.; Lorenzen, S.; Meyer, A.; Randler, C. Lessons with Living Harvest Mice: An empirical study of their effects on intrinsic motivation and knowledge acquisition. *Int. J. Sci. Educ.* **2012**, *34*, 2797–2810. [CrossRef]
58. Duman, İ.; Horzum, M.; Randler, C. Adaptation of the Short Form of the Intrinsic Motivation Inventory to Turkish. *Int. J. Psychol. Educ. Stud.* **2020**, *7*, 26–33. [CrossRef]
59. Hill, L.G.; Betz, D.L. Revisiting the Retrospective Pretest. *Am. J. Eval.* **2005**, *26*, 501–517. [CrossRef]
60. Herzig, B. Wie Wirksam Sind Digitale Medien im Unterricht. 2014. Available online: https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/wie-wirksam-sind-digitale-im-unterricht/ (accessed on 21 June 2022).

61. Müller, H.; Krummenacher, J. Funktionen und Modelle der selektiven Aufmerksamkeit. In *Kognitive Neurowissenschaften*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 307–321.
62. Groß, J.; Langstein, J.; Paul, J.; Ritter, E. Identification of Ants–Development of the Learner-Oriented Digital Tool ID-Logics. *World J. Chem. Educ.* **2020**, *8*, 21–28. [CrossRef]
63. Tessartz, A.; Scheersoi, A. Pflanzen? Wen interessiert's? *Bildungsforschung* **2019**, *1*, 1–22. Available online: https://bildungsforschung.org/ojs/index.php/bildungsforschung/article/view/275/319 (accessed on 15 June 2022).
64. Deci, E.L.; Ryan, R.M. *Handbook of Self-Determination Research*; University Rochester Press: New York, NY, USA, 2004.
65. Colon, J.; Tiernan, N.; Oliphant, S.; Shirajee, A.; Flickinger, J.; Liu, H.; Francisco-Ortega, J.; McCartney, M. Bringing Botany into Focus: Addressing Plant Blindness in Undergraduates Through an Immersive Botanical Experience. *BioScience* **2020**, *70*, 887–900. [CrossRef]
66. Stagg, B.C.; Donkin, M.E.; Smith, A.M. Bryophytes for Beginners: The usability of a printed dichotomous key versus a multi-access computer-based key for bryophyte identification. *J. Biol. Educ.* **2015**, *49*, 274–287. [CrossRef]

sustainability

Article

Predictors of Intention to Use a Sustainable Cloud-Based Quality Management System among Academics in Jordan

Dima Dajani [1], Saad G. Yaseen [2,*], Ihab El Qirem [1] and Hanadi Sa'd [3]

1 Department of Digital Marketing, Faculty of Business, Al-Zaytoonah University of Jordan, P.O. Box 130, Amman 11733, Jordan
2 Department of Business Administration, Faculty of Business, Al-Zaytoonah University of Jordan, Amman 11733, Jordan
3 Faculty of Business, Al-Zaytoonah University of Jordan, Amman 11733, Jordan
* Correspondence: saad.yaseen@zuj.edu.jo

Abstract: This research aims to provide a predictive model of essential factors influencing the behavioral intention to use sustainable cloud-based quality management systems among academics in Jordan. A comprehensive research model was developed based on the Unified Theory of Acceptance and Use of Technology (UTAUT2) and the Theory of Planned Behavior (TPB), which was tested using cross-sectional data. The research sample covers Jordanian higher education institutions (23 governmental and private universities), and the unit of analysis includes 500 academics. The research adapts and modifies the UTAUT2 model and TPB to explain behavioral intention to use sustainable cloud-based quality management systems in developing countries. The proposed model explained 0.478 percent of behavioral intention variance and 0.127 percent of the user behavior variance. Three constructs are found to be significant predictors: perceived behavioral control, performance expectancy, and facilitating conditions. The attitude toward the behavior and subjective norm are not significant predictors. The research contributes to the literature in several ways. First, it extends previous studies by examining predictors of the behavioral intention to use SCQMS in higher education institutions. Second, it provides rigorous empirical evidence that incorporating the UTAUT2 model with the TPB produced a substantial improvement in the variance explained in behavioral intention compared to the prior research conducted in developing contexts. Third, this research provides useful insight into university management. The research provides a better understanding of the essential factors influencing the behavior intention to use sustainable cloud-based quality management systems in Jordanian Universities. Thus, the research model provides better explanatory power than previous studies in business literature and developing markets.

Keywords: behavioral intention; sustainable cloud-based quality management system; theory of planned behavior; unified theory of acceptance and use of technology; IT adoption

Citation: Dajani, D.; Yaseen, S.G.; El Qirem, I.; Sa'd, H. Predictors of Intention to Use a Sustainable Cloud-Based Quality Management System among Academics in Jordan. *Sustainability* **2022**, *14*, 14253. https://doi.org/10.3390/su142114253

Academic Editors: Jian-Hong Ye, Yung-Wei Hao, Yu-Feng Wu, Savvas A. Chatzichristofis and Aras Bozkurt

Received: 6 August 2022
Accepted: 25 October 2022
Published: 1 November 2022

1. Introduction

Technology is essential in a corporate and academic settings. Presently, the quality of education depends on the existence of advanced technologies in higher education institutions. Technology has transferred education from passive and reactive to an interactive and aggressive practice [1], Technological tools assist lecturers in accomplishing work and teaching efficiently. Information technologies develop curiosity and imagination in students' minds and are considered one of the major determinants influencing students' learning intentions [2,3]. Finally, advanced technologies improve the overall quality and integrity of the academic system [4].

Ultimately, the question that concerns policymakers in the higher education sector is how to exploit technological progress to build the capacity of universities in developing countries [1,5]. The value of knowledge and accumulated data are crucial in universities for several decision-makers [6]. The accurate use of information technology affects the long-term sustainability of an organization [7,8]. Thus, universities are seeking out and employing several systems to save and manage their data. Some universities employed sustainable cloud quality management systems to manage their data and process quality information for effective decision-making [9]. Cloud computing is a distributed computing paradigm that provides access to virtual resources such as computers, networks, storage, development platforms, and applications [10]. Thus, it is essential in this research to examine the factors affecting the behavior intention of several staff at universities to use sustainable cloud-based quality management systems.

Previous research concentrated on the conventional quality management system, in which forms, processes, and documentation are completed on paper and manually. Although a sustainable electronic quality management system is efficient and can contribute to an organization's competitive advantage, large capital expenditures associated with traditional local programs, maintenance issues, and updates may be a burden on some educational institutions [11–13]. A solution that helps to alleviate these challenges is to use a sustainable cloud-based quality management system (SCQMS) that allows users to use the same automated system without investing in more expensive equipment or employing and training a significant number of university Information Technology employees [14,15].

Furthermore, while a great deal of literature has been written in the field of paper-based quality systems, there has been little research investigating the essential factors influencing the intention to use SCQMS. Given the particular characteristics of SCQMS, there is a need for additional research to understand the essential factors that affect behavioral intention to accept and use SCQMS. Thus, this research aims to provide a model incorporating essential factors influencing the behavioral intention to use sustainable cloud-based quality management systems in Jordanian universities based on the Unified Theory of Acceptance and Use of Technology (UTAUT2) and the Theory of Planned Behavior (TPB).

This research contributes to the literature in several ways. First, the research extends prior literature by incorporating significant predictors from the UTAUT2 model and the TPB. Second, it fills theoretical gaps by providing empirical evidence on the antecedents of (SCQMS). More specifically, analyzing the intention to use SCQMS in developing countries such as Jordan has not been previously studied. Third, this research provides insights for Jordanian universities' leaders about how academics engage in pursuing cloud computing quality management systems. Forth, the current research sheds light on the most significant predictors: namely perceived behavioral control. Finally, our research provides a better understanding of the critical factors influencing the intention to use SCQMS in Jordanian universities Thus, the research intends to address this issue and propose a model for assessing the behavioral intention to use SCQMS at Jordanian universities. The emphasis on Jordanian universities is appropriate given that the majority of research has focused on SCQMS in developed countries.

2. Literature Review

2.1. Quality Management System (QMS)

Colleges and universities strive to increase their competitiveness by finalizing tasks efficiently and effectively [16]. Information is an essential component of the traditional quality management system. The primary purpose of QMS is to give appropriate process execution and process standards [17,18]. As the volume of information grows, various software are required to produce, gather, store, process, and communicate the information required to complete fundamental activities and objectives [18].

QMS is a collection of processes, papers, policies, objectives, and manuals that offer a framework for the universities operation to fulfill the demands of academic staff, continually develop the organization, and use standardized techniques. Chen and Wu [5] defined a quality management system as a model, method, and tool for achieving business objectives [19].

In higher education institutions, QMS refers to integrating the quality management system and applying and implementing quality management technology tools within the universities to improve organizational performance [20,21]. The successful adoption of QMS requires universities to comply with a set of standards, including a series of specific rules, regulations, and policies to regulate university activities and standardizing methods to guide existing quality processes [16,22–24].

At the moment, businesses, educational institutions, and governments are working together to create a new platform for increasing mobile Internet service quality [25]. As a result, cloud computing has become a significant milestone in the development of information systems and quality management [26].

2.2. Sustainable Cloud-Based Quality Management System

Cloud computing and quality management systems are unavoidable trends in computer science, information systems, and quality management systems [27]. Cloud computing is an emerging innovation phenomenon in information technology [28]. Cloud computing enabled direct interaction with data centers without the need for additional hardware or software, resulting in better sharing and use of resources and costs [29]. Sustainability seeks to prevent the depletion of natural or physical resources so that they remain available in the long-term and contribute to maintaining or supporting the process that continues over time [30,31].

Designing a sustainable system is one of the biggest challenges of the 21st century: environmental transformation combined with digital transformation [32]. SCQMS is a subset of cloud computing in the field of quality management systems [33]. SCQMS has all of the materials and equipment, such as hardware and software resources, to improve the traditional QMS infrastructure. Once the set of SCQMS documents and models are installed by default on cloud servers, these documents and forms are ready for use by academic staff at universities from cloud vendors [34–36].

SCQMS assists universities in reducing capital costs, maintaining issues, and training, allowing them to focus more on their primary activities (teaching) rather than managing issues in information technologies that will be delivered to them by the service provider. Furthermore, using a sustainable cloud-based quality management system minimizes energy usage and maintains green and environment-friendly universities. Furthermore, SCQMS assists academic staff in organizing, categorizing, reporting, managing, and accessing their daily work information at any time and from any location. As a result, it improves the educational process and learning outcomes [10,25,37–41]. Therefore, employing a sustainable cloud-based quality management system provides a significant edge over rivals who are still using paper-based quality management systems or a variety of offline tools.

2.3. The TPB and the UTAUT2

The theory of planned behavior (TPB) is an extension of the theory of reasoned action [42,43]. A central fact in the theory of planned behavior is the individual's intention to perform a specific behavior the stronger the intention to perform a given behavior, the more likely should be its performance. According to the theory of planned behavior, perceived behavior control together with behavioral intention can be used directly to predict behavior (performance, action). Furthermore, the theory of planned behavior postulates three conceptual determinants of behavioral intention. The first is the attitude toward the behavior in a question. The second is a social factor, termed subjective norm, and the third antecedent of the behavioral intention is the degree of behavioral control [44].

Attitude toward the behavior refers to the degree to which a person has a favorable or unfavorable evaluation of the behavior in a question. Subjective norms refer to the social pressure to perform or not perform the behavior, whereas perceived behavioral control describes the perceived ease or difficulty of performing the behavior; it is assumed to reflect experience, as well as anticipated impediments and obstacles.

In combination, attitude, subjective norm, and perception of behavioral control led to the formation of a behavioral intention. The intention is thus assumed to be the immediate antecedent of behavior.

Venkatesh et al. [45] reviewed IT users' acceptance literature and discussed eight prominent IT acceptance/adoption theoretical models. The eight models make up the social cognitive theory. Using data from four organizations over six months with three points of measurement, the eight models explained between 17% and 53% of the variance in user intentions [45].

Based on these results, Venkatesh et al. [45] developed a new and integrated model called the "Unified Theory of Acceptance and Use of Technology" within four essential constructs and four moderators. The new model, UTAUT, was then tested using the original data and found to outperform the eight theoretical models.

The model shows that four determinants have a significant influence on the user's intention to accept and use IT systems, namely: performance expectancy, effort expectancy, social influence, and facilitating conditions.

Moreover, Venkatesh et al. [46] extended the theoretical framework model (UTAUT) to study IT acceptance and use in a consumer setting. The new proposed model is called UTAUT2. The UTAUT2 model incorporates three new constructs into UTAUT: hedonic motivation, price value, and habit. Individual differences–namely age, gender, and experience–are hypothesized to moderate the influence of these predictors on the user's intention to accept or use a specific technology. The extended UTAUT2 model produced a substantial improvement in the variance explained in behavioral intention (56% to 74%) compared to the UTAUT model [47].

Hedonic motivation (conceptualized as perceived enjoyment) can be defined as the fun derived from using a specific technology. In the business literature, hedonic motivation has been found to influence users' intention to accept technology [47,48].

Price value has a significant influence on user's technology acceptance, especially in the e-service marketing context [47,49], whereas habit has been used as the extent to which the user tended to perform behaviors automatically, in terms of the learning curve accumulation [50,51].

3. Research Model and Hypothesis

Davis [52] proposed the Technology Acceptance Model (TAM) based on the Theory of Reasoned Action (TRA). TAM aims to identify factors that explain the acceptance and use of various information technologies and to indicate the adjustments that need to be made in the system to make it user-friendly [52,53].

The Theory of Planned Behavior (TPB) was developed by Icek Ajzen to predict human behavior [44]. The TPB is an extension of the Theory of Reasoned Action (TRA) [42,43]. This theory focused primarily on the intention of a person to perform a particular behavior; the intention to perform a behavior is stronger when it is likely to perform. TPB assumes that attitude toward the behavior, subjective norm, and perceived behavioral control influence behavioral intention.

Venkatesh et al. [45] examined and discussed eight IT user acceptance/adoption models. The eight models were TRA, TAM, TPB, the motivational model, the innovation diffusion theory, the social cognitive theory, a model combining TAM and TRA, and the model of PC use. Venkatesh et al. [45] developed The Unified Theory of Acceptance and Use of Technology (UTAUT) using the suggested complete synthesis of these eight leading models. There are four core constructs in UTAUT: performance expectancy (PE), effort expectancy (EE), social influencing (SI), and facilitating conditions (FC). Several studies on the adoption of different IT system applications [54–56] showed that UTAUT constructs are generalizable. However, Venkatesh et al. [47] developed the UTAUT2 model and expanded it to explain the use of technology in the consumer context. The purpose of this model is to explain the user's intention to use various information systems. The UTAUT2 model incorporates three new constructs into UTAUT: hedonic motivation, price value, and habit [47].

This research adopted the UTAUT2 model because of its comprehensiveness, and improved predictive power which exceeded many technologies acceptance models, such as UTAUT and TAM. Models that focus on identifying essential predictors and determinants are considered to be vital in providing a rich understanding of the user's acceptance and use of technology. Venkatesh [47] asserted that in the case of UTAUT, which was originally developed to explain user technology acceptance and use, it will be critical to examine how it can be extended to other contexts, such as the context of the sustainable cloud-based quality management system acceptance or adoption.

The research model incorporated two constructs from the UTAUT2 which are performance expectancy and facilitating conditions. The variables of hedonic motivation (which is defined as the fun derived from using a specific technology), price value, and habit were excluded because they have a direct bearing on consumer behavior and have no effect on this study because this technology is adopted by enterprises that invest in the system. In addition, the use of a sustainable cloud-based quality management system is compulsory among academics in Jordanian universities. Therefore, hedonic motivation is more suitable to incorporate in voluntary settings.

Furthermore, these variables are not suitable for the current study that is dealing with the topic of sustainable cloud-based quality management systems among academics in Jordanian universities.

In addition, the research model adapts three constructs from the TPB model (attitude, subjective norm, and perceived behavioral control) that influence behavioral intention to use a sustainable cloud-based quality management system, as shown in Figure 1.

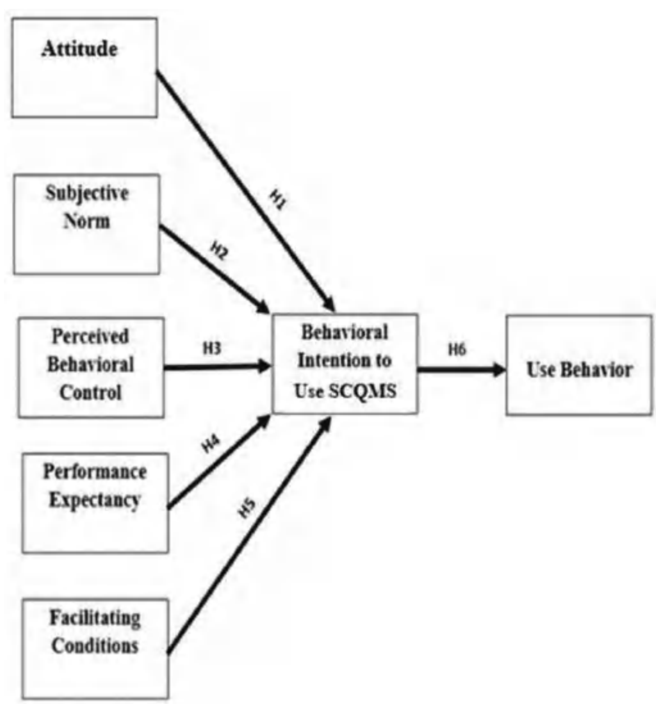


Figure 1. Conceptual Research Model Adapted from Venkatesh et al. [47] and Ajzen [44].

The extended unified theory of acceptance and use of technology, presented by Venkatesh et al. [46], was used in different contexts in the Middle East and the Arab world. For example, Dajani and Hegleh [57] used extended UTAUT2 to test the antecedents that impact on behavior intention of animation usage among marketing students in Jordanian universities. Abu Shakra and Nikbin [58] used UTAUT2 to explore and discuss the factors that impact the acceptance and adoption of the Internet of Things by entrepreneurs in Oman. El-Masri and Tarhini [59] used UTAUT2 to examine the main factors that may impede or enable the adoption of e-learning systems by university students from developing countries (Qatar) and developed countries (the United States).

However, the extended UTAUT2 had never been used to explain the use of SCQMS among academics in Jordanian universities. Cloud computing was investigated in Jordanian universities, such as in the study by Matar et al. [29] but this did not focus on sustainable quality management. Other cloud computing technologies were already assessed in Jordan, such as the cloud-based accounting information system adoption by Alshirah et al. [60], and cloud computing adoption in hospitals by Harfoush et al. [28]. These prior works are not transferable to the current research-specific cloud computing application. Therefore, the current research adopted the TPB and UTAUT2 models, which are primarily concerned with an individual's intention to perform a certain behavior and explain key aspects impacting the behavioral intention to use a sustainable cloud-based quality management system in Jordanian universities.

3.1. Attitude

Attitude refers to the extent to which a person makes a positive or negative judgment of actions [44]. In this context, attitude reflects the degree to which an academic staff member has an adequate or inappropriate assessment of their ability to employ SCQMS. The Theory of Planned Behavior states that attitude toward a behavior is a good predictor of intention and behavior. Several models and hypotheses have been proposed by researchers to investigate the relationship between attitude and behavior [61]. Wicker [62] investigated the relationship between attitude and behavior and discovered that the two may not be related. Hale et al. [63] stated that there is a limited association between attitude and behavior in a voluntary setting due to dissatisfaction with previous experience. Furthermore, research demonstrated that behavioral intention to use SCQMS is positively influenced by one's attitude toward the behavior [44,64].

Therefore, the following hypothesis is proposed:

H1: *Attitude positively impacts the behavioral intention to use SCQMS.*

3.2. Subjective Norm

Subjective Norma refers to social pressure to perform or refrain from performing a behavior [44]. According to the empirical data, some studies discover favorable impacts while others find negative consequences. Personal concerns, according to Ajzen, I [44], may lessen the subjective norm impact. According to Shon et al. [40], the subjective norm is the weakest variable in TPB and has less predictive power in forecasting people's intents. According to this assertion, some researchers deleted subjective norms from their studies and replaced them with personal normative beliefs [40,65–67].

On the other hand, other researchers argue that subjective norms are still required for predicting individuals' intentions [68–71]. When Armitage and Conner [72] used multiple-component measurements, they discovered a substantial link between subjective norm and intention. Additionally, the subjective norm has been proven to be an enabler for behavioral intention measurements [73]. As a result, it is proposed that subjective norm positively influences behavioral intention to use SCQMs.

H2: *Subjective norm positively impacts the behavioral intention to use SCQMS.*

3.3. Perceived Behavioral Control

Perceived behavioral control is an individual's assessment of having the internal capability to exert external control over a specific behavior [74]. In the framework of SCQMS, academic staff considers the degree of control required to complete an activity.

It has been demonstrated that perceived behavioral control is a facilitator for behavioral intention measurements [75]. As a result, it is postulated that perceived behavioral control influences behavioral intention to use SCQMS. The resulting proposition is suggested:

H3: *Perceived behavioral control positively impacts the behavioral intention to use SCQMS.*

3.4. Performance Expectancy (PE)

PE is defined as the degree to which people feel that adopting the system would assist them in work performance [47]. When compared to other particular combinations in the UTAUT models to employ new technology, PE was determined to be the most potent driver of behavior intention [27,45,47,56,57,76]. According to current literature, people prefer to accept novel technologies when they are of value to them [76]. The view of academic staff that the use of SCQMS will be beneficial in their administrative and teaching duties is referred to as a performance expectation. As a result, this research proposes that performance expectancy influences behavioral intention to use SCQMS. As a result, the following hypothesis is developed:

H4: *Performance Expectancy has a positive impact on the behavioral intention to use SCQMS.*

3.5. Facilitating Conditions

Facilitating Conditions refer to the degree of accessibility to the tools and assets required to complete a task [47]. Academic staff performs their tasks smoothly when enough resources are provided. Facilitating conditions have been verified as a significant guide for SCQMS adoption and use [27,45,47,56,57]. Facilitating conditions include everything that aids in the implementation of the assessment method, such as administrative, organizational, or technical assistance, expertise, and other resources [76]. In the context of academia, facilitating conditions refer primarily to the adequate technological infrastructure (e.g., internet access) and the availability of a technical expert willing to assist academic staff in overcoming technical issues that may develop while using SCQMS. As a result, this research proposes the following:

H5: *Facilitating conditions have a positive impact on the behavioral intention to use SCQMS.*

3.6. Behavioral Intention

Behavioral intention is described as a measure of the intensity of one's desire to engage in a given action [43]. This construct describes the strength of academic staff to use SCQMS. The importance of behavioral intention as an indicator of individual behavior is well documented in IT literature [45,56,64,77,78]. According to Fishbein and Ajzen [42], people's intentions affect their behavior and action. According to Venkatesh [45,47] behavioral intention has a significant influence on technology use. As a result, the following hypothesis is proposed:

H6: *Behavioral intention has a positive impact on the use of SCQMS.*

4. Methods

4.1. Sampling and Data Collection

This research aims to investigate the behavioral intention to use sustainable cloud-based quality management systems in Jordanian universities. Several research hypotheses have been proposed to examine the extent to which essential factors (attitude, subjective norm, perceived behavioral control, performance expectancy, and facilitating conditions) can explain behavioral intention to use SCQMS in Jordanian universities. After developing measurement scales an instrument was designed in the form of a self-administered questionnaire using a five-point Likert scale. The questionnaire was distributed online to encourage respondents to fill in the survey questionnaire. A total of 568 questionnaires were distributed online to academics. Of those, 500 were received for a gross response rate of 88%. The target population of this research covers Jordanian universities (11 government universities and 12 private universities). The sample for this research comprised 23 Jordanian universities. The research unit of analysis includes 500 academics, of whom 88 are professors, 164 associate professors, 208 assistant professors, and 40 lecturers who were randomly selected. Academics were selected because they are the primary users of SCQMS in Jordanian universities.

Among the respondents, 83.6% were male and 16.4% were female. The majority of the respondents were between 40 and 49 years old and constitute 40% of the sample. Meanwhile, 35.2% of the respondents have 5–9 years of experience.

Table 1 describes the demographic profile of the respondents.

Table 1. Demographic profile of respondents.

	Frequency (N—500)	Percentage (%)
Academic Rank		
Professor	88	17.6%
Associate Professor	164	32.8%
Assistant Professor	208	41.6%
Lecturer	40	8%
Total	**500**	**100%**
Gender		
Male	418	83.6%
Female	82	16.4%
Total	500	100%
Age		
less than 30	12	2.4%
30–39	136	27.2%
40–49	202	40.4%
51-above	150	30.0%
Total	500	100%
Years of experience		
less than 5	78	15.6%
5–9	176	35.2%
10–14	108	21.6%
15-above	138	27.6%
Total	500	100%

4.2. Measurement

Most of the measurements were adapted from prior research. Cheng et al.'s [79] scale was used to assess the attitude variable. The subjective norm was measured using the scale in [80]. Perceived behavioral control was measured using the items in [81]. Performance expectancy was measured using [45,47]. Facilitating conditions were also measured using [45,47]. Venkatesh et al.'s [45,47] scale was used to assess behavior intention and user behavior. The optimization of the research constructs is illustrated in Table 2.

Quantitative survey research appeared to be appropriate to collect the research data and examine the model. All research constructs were measured with a five-point Likert scale, and they were formulated using positive statements. The research survey questionnaire was created in English and then translated into Arabic. Back translation was used to ensure correctness.

Table 2. Constructs Measurements.

Construct	Code	Measurements	References
Attitude	ATT1	Using a sustainable cloud-based quality management system would be a good idea	Cheng et al. [79]
	ATT2	Using Use sustainable cloud-based quality management system would be a foolish idea	
	ATT3	I like the idea of using the sustainable cloud-based quality management system	
	ATT4	Using a sustainable cloud-based quality management system would be pleasant	
Subjective Norm	SN1	Your decision to use a sustainable cloud-based quality management system is because universities use this system	Madden et al. [80]
	SN2	Your decision to use a sustainable cloud-based quality management system is because the media encourages s use of this system	
	SN3	Your decision to use a sustainable cloud-based quality management system is because International Higher educational institutions use this system	
Perceived Behavioral Control	PBC1	I have control over using the sustainable cloud-based quality management system	Wu and Chen [81]
	PBC2	I have the resources necessary to use a sustainable cloud-based quality management system	
	PBC3	I know it is necessary to use a sustainable cloud-based quality management system	
	PBC4	Given the resource, opportunity, and knowledge it takes to use a sustainable cloud-based quality management system, it would be easy for me to use SCQMS	
Performance Expectancy	PE1	I find using a sustainable cloud-based quality management system useful in my daily work	Venkatesh et al. [45,47]
	PE2	Using a sustainable cloud-based quality management system increases my chances of achieving things that are important to me	
	PE3	Using a sustainable cloud-based quality management system helps me accomplish things more quickly	
	PE4	Using a sustainable cloud-based quality management system increases my productivity	
Facilitating Conditions	FC1	I have the resources necessary to use a sustainable cloud-based quality management system	
	FC2	I know that it is necessary to use a sustainable cloud-based quality management system	
	FC3	Using a sustainable cloud-based quality management system is compatible with other technologies I use	
	FC4	I can get help from others when I have difficulties using the sustainable cloud-based quality management system	

Table 2. *Cont.*

Construct	Code	Measurements	References
Behavior Intention	BI1	I intend to continue using sustainable cloud-based quality management system in the future	Venkatesh et al. [45,47]
	BI2	I will always try to use a sustainable cloud-based quality management system in my daily work	
	BI3	I plan to continue to use sustainable cloud-based quality management systems frequently	
User Behavior	UB1	If an initial decision to use SCQMS has been taken, how frequently will you use it?	Venkatesh et al. [45,47]
	UB2	If an initial decision is to use SCQMS, how much time will you spend on it in terms of minutes/hours?	

4.3. Data Analysis Results

4.3.1. The Measurement Model

For data analysis, partial least squares structural equation modeling (PLS-SEM) version 3.0 was employed. Smart PLS is useful for exploratory research with a small sample. Partial least squares equation modeling (SEM) has been used in a variety of business research fields [82]. A PLS path model also has two components: the measurement model (inner model) and the structural model (outer model). The measurement model offered data about the scales' reliability and validity, while the structured model reflected the linkages (paths) between the research components [83].

The validation of the measurement model is determined by the assessment of convergent and discriminant validity. All items loading, Cronbach's alpha, composite reliability, and average variance extracted exceeded the commonly accepted thresholds (0.7), implying that all items consistently represent and measure the same construct [84–87].

Discriminate validity refers to the distinctiveness of the research's model construct using three different techniques: Fornell and larker criterion, cross-loadings, and Heterotrait-Monotrait (HTMT) ratio. Table 3 shows the Fornell-Larker criterion and Table 4 illustrates cross-loadings. As seen in Tables 4 and 5, the square root of the average variance extracted (AVE) of each latent construct is bigger than its strongest association with other constructs.

Table 3. Fornell and Larker's criterion.

	Attitude	Behavioral Intention	Facilitating Condition	Perceived Behavioral Control	Performance Expectancy	Subjective Norm	Use Behavior
Attitude%	0.869						
Behavioral Intention%	0.317	0.891					
Facilitating Condition%	0.180	0.408	0.859				
Perceived Behavioral Control%	0.312	0.435	0.718	0.793			
Performance Expectancy%	0.399	0.649	0.328	0.350	0.842		
Subjective Norm%	0.577	0.374	0.302	0.351	0.461	0.825	
Use Behavior%	0.194	0.356	0.220	0.255	0.336	0.275	0.901

Table 4. Cross-Loading.

Constructs	Attitude	Subjective Norm	Perceived Behavioral Control	Performance Expectancy	Facilitating Condition	Behavioral Intention	Use Behavior
ATT1r	**0.873**	0.492	0.288	0.374	0.180	0.241	0.201
ATT2r	**0.869**	0.525	0.278	0.387	0.190	0.263	0.224
ATT3r	**0.893**	0.482	0.240	0.313	0.081	0.308	0.149
ATT4r	**0.840**	0.511	0.285	0.322	0.188	0.280	0.110
SN1r	0.440	**0.825**	0.352	0.428	0.320	0.298	0.206
SN2r	0.539	**0.841**	0.311	0.349	0.253	0.351	0.277
SN3r	0.437	**0.809**	0.192	0.370	0.165	0.265	0.185
PBC1r	0.312	0.325	**0.774**	0.253	0.475	0.302	0.207
PBC2r	0.262	0.284	**0.844**	0.248	0.694	0.312	0.207
PBC3r	0.172	0.245	**0.789**	0.286	0.680	0.323	0.218
PBC4r	0.246	0.263	**0.764**	0.309	0.451	0.416	0.182
PE1r	0.388	0.385	0.386	**0.842**	0.346	0.585	0.282
PE2r	0.287	0.352	0.260	**0.840**	0.257	0.512	0.273
PE3r	0.349	0.427	0.286	**0.845**	0.270	0.550	0.341
PE4r	0.312	0.386	0.238	**0.843**	0.224	0.535	0.236
FC1r	0.169	0.233	0.687	0.219	**0.831**	0.289	0.130
FC2r	0.187	0.284	0.633	0.316	**0.880**	0.404	0.204
FC3r	0.105	0.254	0.543	0.296	**0.865**	0.344	0.222
BI1r	0.314	0.334	0.377	0.641	0.324	**0.896**	0.335
BI2r	0.211	0.308	0.385	0.523	0.379	**0.874**	0.263
BI3r	0.313	0.357	0.404	0.565	0.393	**0.903**	0.348
UB1r	0.250	0.287	0.292	0.369	0.237	0.377	**0.943**
UB2r	0.063	0.191	0.139	0.208	0.141	0.241	**0.856**

Table 5. The measurement models.

Construct	Code	Loading	Cronbach's Alpha	Composite Reliability	Average Variance Extracted (AVE)
Attitude	ATT1 ATT2 ATT3 ATT4	0.873 0.869 0.893 0.840	0.892	0.925	0.755
&Subjective Norm	SN1 SN2 SN3	0.825 0.841 0.809	0.768	0.865	0.681
Perceived Behavioral Control	PBC1 PBC2 PBC3 PBC4	0.774 0.844 0.789 0.764	0.805	0.872	0.629
Performance Expectancy	PE1 PE2 PE3 PE4	0.842 0.840 0.845 0.843	0.864	0.907	0.709
Facilitating Condition	FC1 FC2 FC3	0.831 0.880 0.865	0.824	0.894	0.738
Behavioral Intention	BI1w BI2w BI3w	0.896 0.874 0.903	0.871	0.920	0.794
Use Behavior	UB1 UB2	0.943 0.856	0.777	0.895	0.811

Regarding cross-loading, a particular item of each construct should have higher loading on its construct in comparison to other constructs. As Table 4 shows, the value of factor loading for each item also satisfies the requirement of (0.70).

As Table 5 shows, the composite reliability values range from 0.865 to 0.925, with an average value extracted of more than 50%. AVE is a typical measure used to prove the convergent validity of the construct. An AVE value of 0.50 or more suggests that the concept explains more than half of the variation in its indicators [83].

4.3.2. The Structural Model

Cross-validation (CV)-communality redundancy indices, path coefficients (β), and coefficient of determinator (R^2) values are important measurements for evaluating the research reflective structural model. Cross-validation and the blindfolding-based cross-validated redundancy measure and the statistical significance of the path coefficients are used to evaluate the structural model's quality. For all constructs, this index should be positive [83].

Furthermore, estimates for the hypothesized relationships between the constructs are produced after running the PLS-SEM method. Figure 2 depicts the results of the structural model research, including predicted path coefficients (β) and the R^2 value of the determination coefficient. R^2 explains the variance in the endogenous constructs explained by exogenous constructs [88]. Furthermore, Henseler et al. [89] suggest that an HTMT value above 0.90 depicts a lack of discriminant validity. Overall, the research model accounts for 0.478 of the variances in behavioral intention and 0.127 of the variances in user behavior, as shown in Figure 2.

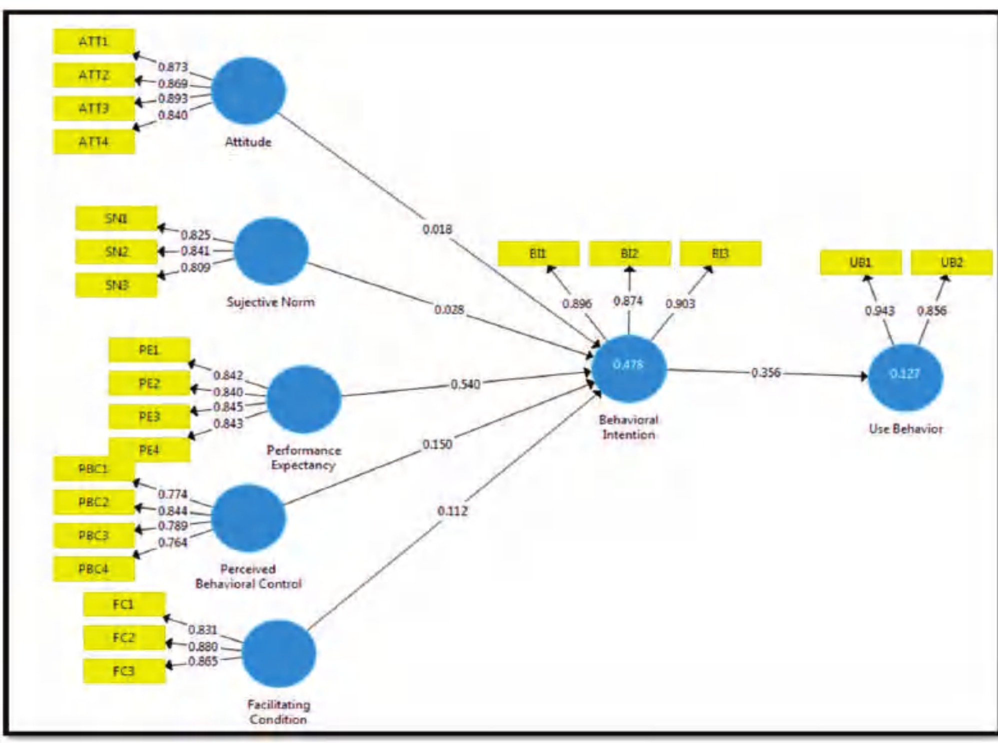


Figure 2. The Structural Model Results.

Table 6 presents the outcomes of the hypotheses that were tested. The results indicate a strong positive effect for four hypotheses H3, H4, H5, and H6, while H1 and H2 are not supported. Consequently, H1 the results for attitude are (β = 0.018), T-Value (0.435) and $p > 0.05$. This result refers to the rejection of the proposition. Therefore, H1 is not supported, and attitude has no positive impact on behavioral intention to use SCQMS. H2 the results for subjective norm are (β = 0.028), T-Value (0.698) and $p > 0.05$. This result refers to the rejection of the proposition. Therefore, H2 is not supported, and the subjective norm has no positive impact on behavioral intention to use SCQMS. H3 the path coefficient for perceived behavioral control is (β = 0.150) and has a value of (T-Value = 3.027) at a significant level of ($p < 0.05$). This outcome refers to the acceptance of the hypothesis. It could be concluded that there is a positive effect of perceived behavioral control on behavioral intention to use SCQMS. H4 the path coefficient for performance expectancy is (β = 0.540) and has a value of (T-Value = 15.904) at a significant level of ($p < 0.05$). This outcome refers to the acceptance of the hypothesis. It could be concluded that there is a positive effect of performance expectancy on behavioral intention to use SCQMS. H5 the path coefficient for facilitating condition is (β = 0.112) and has a value of (T-Value = 2.003) at a significant level of ($p < 0.05$). This outcome refers to the acceptance of the hypothesis. It could be concluded that there is a positive effect of facilitating conditions on behavioral intention to use SCQMS. H6 the path coefficient for behavioral intention is (β = 0.356) and has a value of (T-Value = 8.944) at a significant level of ($p < 0.05$). This outcome refers to the acceptance of the hypothesis. It could be concluded that there is a positive effect of behavioral intention on use behavior.

Table 6. The structural model results.

Path	Standardized Coefficient Beta (β)	T-Values	P-Values (Sig)	Hypothesis Results
Attitude -> Behavioral Intention3	0.018	0.435	0.664	H1 Not Supported1
Subjective Norm -> Behavioral Intention4	0.028	0.698	0.486	H2 Not Supported2
Perceived Behavioral Control -> Behavioral Intention5	0.150	3.027	0.003	H3 Supported
Performance Expectancy -> Behavioral Intention	0.540	15.904	0.000	H4 Supported
Facilitating Condition -> Behavioral Intention	0.112	2.003	0.046	H5 Supported
Behavioral Intention -> Use Behavior	0.356	8.944	0.000	H6 Supported

5. Discussion and Implications and Limitations

The research investigated the antecedents that affect the academic staff's behavioral intention to use SCQMS by adapting the TPB and UTAUT2 models. The proposed structural model is acceptable and the relevance constructs have a strong ability to demonstrate the various factors that influence the behavioral intention of academic staff to use the SCQMS.

Empirically, the results of the analyses indicated that the link between performance expectancy and the behavioral intention to use the SCQMS is the most significant. In addition, academic staff at universities find that using the SCQMS is useful in carrying out their daily work as it increases the chances of achieving things that matter to them and helps them to complete their work more quickly, therefore increasing their productivity. This result is in agreement with [26,27,45,47,56,57,90,91].

The construct of facilitating conditions also has an important influence on the behavioral intention of academic staff to use the SCQMS. Academic staff feel that universities have the resources they need to use SCQM, which is compatible with other techniques used in their work. Academic staff also feel that they have the knowledge and ability to use SCQMS and that if they encounter difficulties, they can obtain help from others. These results are consistent with Venkatesh et al. [90], Williams et al. [56]; Venkatesh et al. [45]; Venkatesh et al. [47]; Sharif et al. [91]; Nguyen et al. [27].

Moreover, perceived behavioral control has a significant influence on the behavioral intention to use the SCQMS. The academic staff feel the system is user-friendly because of the resources, opportunities, and knowledge required. This result is in agreement with Yen et al. [92], Ajzen [44], and Jafarkarimi et al. [93].

The hypothesized association between behavioral intention to use SCQMS and use behavior was significant. It shows how frequently the SCQMS system is used and how long academic staff spends using the system [27,38,44,45,47,56].

In addition, the outcome of this research analysis showed that attitude did not have any influence on the behavior intention to use SCQMS. Moreover, academic staff did not regard attitudes as having an impact on their behavior. The attitude failed to predict intention, which means that there was a behavioral gap, which was recognized by many researchers [94,95]. Perhaps attitude does not have an effect because staff are obliged to use the system at their universities.

Furthermore, the subjective norm does not affect the behavior intention to use SCQMS. The decision of academic staff to use the SCQMS did not arise based on social pressure (other universities, media, and international higher educational institutions). Therefore, the subjective norm has no significant influence on the behavioral intention to use the SCQMS. The related outcome was also found by Sheppard et al. [96], Davis et al. [97], Mathieson [98], and Yen et al. [92]. The research findings did not reveal the significance of the attitude and subjective norm on the intention to adopt SCQMS.

The research contributes to the literature because it has a different contextual setting in comparison with the traditional technology acceptance models which are predominantly Western in origin, it incorporates different constructs that are vital to explain Jordanian and developing economy issues and it illustrates a different methodology for operationalization of the constructs and testing of the hypotheses. Therefore, the study contributes to the literature concerning technology acceptance and use in developing countries and higher education sector.

On the theoretical side, the current research contributes to the critical review of literature on the reliability and validity of the well-known model UTAUT2 and its ability to predict behavioral intention to adopt and use a specific digital system in a developing country context. The research model incorporates constructs from various technology models to explain the behavior intention to use the system. Thus, the authors argue that incorporating predictors from other IT adoption models is essential to enhance our understanding of innovative technologies and human behavior. Finally, the investigated technology, namely the sustainable cloud-based quality management system among academics has not been investigated in Jordanian universities. Therefore, this study adds value to the literature in developing countries and more specifically in the educational sector.

The present research sheds light on the essential role of perceived behavior control, performance expectancy, and facilitating conditions in predicting intention to adopt and use SCQMS. The study indicates that the technology characteristics, such as performance expectancy and compatibility, can increase the acceptance and use of SCQMS. Therefore, programmers and designers of SCMS should pay attention to the usefulness, the ease of use and compatibility of the system. It is recommended to create systems that are easy to use, useful, interactive, and compatible with the needs of the academics to help them understand and allocate what they are searching for. Furthermore, the language and the instruction of the system should be easy to understand. Software programmers should develop software that has a bi-lingual interface (Arabic and English) to be used and understood by all of the employees at Arab organizations.

Finally, decision-makers and IT specialists at universities should provide relevant and sufficient training for all the academics to increase their familiarity with the system and consequently their productivity. IT specialists should be always available to help academics solve any issues related to the use of the system. Finally, responsible parties at universities should allow access for all the employees to use the system after excessive training to facilitate knowledge sharing among all employees.

6. Conclusions and Future Research

The current research results show that three of the independent variables are good predictors of the dependent variable construct (behavioral intention to use a sustainable cloud-based quality management system in the research model). The independent variables are perceived behavioral control, performance expectancy, and facilitating conditions.

Furthermore, the results present the participation of each main predictor in the behavioral intention to use SCQMS. The largest Beta value in the standardized coefficient is (0.540), which returns to the performance expectancy predictor. This variable helps explain behavior intentions with a sustainable cloud based on a quality management system. This is followed by perceived behavioral control (0.150) and then facilitating condition (0.112).

In contrast, the attitude and subjective norm are not significant predictors in the model. The results of the relationship between behavioral intention to use CQMS and the use behavior indicate a significant between the two variables. Behavioral intention contributes (0.356) to explaining the dependent variable variance.

Similar to any empirical study, this research has a few limitations which can be assumed as new directions for future research. Future research could examine other predictors and other mediated constructs such as perceived trust, user satisfaction, or cultural values. Furthermore, the cross-sectional research design is incapable of confirming the predictive power of the causal relations empirically. Future research could address this issue using a comparative and longitudinal research design.

Author Contributions: Formal analysis, D.D.; Methodology, D.D.; Writing—review & editing, D.D.; Data curation, D.D.; Writing—original draft, S.G.Y.; Investigation, S.G.Y.; Supervision, S.G.Y.; Conceptualization, I.E.Q.; Project administration, I.E.Q.; Resources, H.S.; Visualization, H.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The manuscript is conformed with MDPI research data policies.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Al-Samarraie, H.; Saeed, N. A systematic review of cloud computing tools for collaborative learning: Opportunities and challenges to the blended-learning environment. *Comput. Educ.* **2018**, *124*, 77–91. [CrossRef]
2. Chen, C.H.; Tsai, C.C. In-service teachers' conceptions of mobile technology-integrated instruction: Tendency towards student-centered learning. *Comput. Educ.* **2021**, *170*, 104224. [CrossRef]
3. Dutta, B.; Peng, M.H.; Chen, C.C.; Sun, S.L. Interpreting Usability Factors Predicting Sustainable Adoption of Cloud-Based E-Learning Environment during COVID-19 Pandemic. *Sustainability* **2021**, *13*, 9329. [CrossRef]
4. Amigud, A.; Arnedo-Moreno, J.; Daradoumis, T.; Guerrero-Roldan, A.E. An integrative review of security and integrity strategies in an academic environment: Current understanding and emerging perspectives. *Comput. Secur.* **2018**, *76*, 50–70. [CrossRef]
5. Chen, C.L.; Wu, C.C. Student's behavioral intention to use and achievements in ICT-Integrated mathematics remedial instruction: A case study of a calculus course. *Comput. Educ.* **2020**, *145*, 103740. [CrossRef]
6. Giebels, D.; van Buuren, A.; Edelenbos, J. Using knowledge in a complex decision-making process–Evidence and principles from the Danish Housing project's ecosystem-based management approach. *Environ. Sci. Policy* **2015**, *47*, 53–67. [CrossRef]
7. Al Omoush, K.S.; Al-Qirem, R.M.; Al Hawatmah, Z.M. The degree of e-business entrepreneurship and long-term sustainability: An institutional perspective. *Inf. Syst. e-Bus. Manag.* **2018**, *16*, 29–56. [CrossRef]
8. Al-Shayea, Q.K. Telemedicine using cloud computing in Jordan. In *Competition Forum*; American Society for Competitiveness: Tysons Corner, VA, USA, 2015; Volume 13, No. 2; p. 390.
9. Pańkowska, M.; Pyszny, K.; Strzelecki, A. Users' adoption of sustainable cloud computing solutions. *Sustainability* **2020**, *12*, 9930. [CrossRef]
10. González-Martínez, J.A.; Bote-Lorenzo, M.L.; Gómez-Sánchez, E.; Cano-Parra, R. Cloud computing and education: A state-of-the-art survey. *Comput. Educ.* **2015**, *80*, 132–151. [CrossRef]
11. Chen, C.-S.; Liang, W.-Y.; Hsu, H.-Y. A cloud computing platform for ERP applications. *Appl. Soft Comput.* **2015**, *27*, 127–136. [CrossRef]

12. López, C.; Ishizaka, A. GAHPSort: A new group multi-criteria decision method for sorting a large number of the use sustainable cloud-based ERP solutions. *Comput. Ind.* **2017**, *92*, 12–25. [CrossRef]
13. Cheng, Y.M. Quality antecedents and performance outcome of use sustainable cloud-based hospital information system continuance intention. *J. Enterp. Inf. Manag.* **2020**, *33*, 654–683. [CrossRef]
14. Hew, T.-S.; Kadir, S.L.S.A. Predicting instructional effectiveness of use sustainable cloud-based virtual learning environment. *Ind. Manag. Data Syst.* **2016**, *116*, 1557–1584. [CrossRef]
15. Wang, J.; Jou, M. Qualitative investigation on the views of inquiry teaching based upon the cloud learning environment of high school physics teachers from Beijing, Taipei, and Chicago. *Comput. Hum. Behav.* **2016**, *60*, 212–222. [CrossRef]
16. Díaz, J.A.A.; Martínez-Mediano, C. The impact of ISO quality management systems on primary and secondary schools in Spain. *Qual. Assur. Educ.* **2018**, *26*, 2–24. [CrossRef]
17. Grudzień, Ł.; Hamrol, A. Information quality in design process documentation of quality management systems. *Int. J. Inf. Manag.* **2016**, *36*, 599–606. [CrossRef]
18. Laudon, K.C.; Laudon, J.P. *Management Information Systems*; Pearson: Upper Saddle River, NJ, USA, 2015; p. 143.
19. Rocha-Lona, L.; Garza-Reyes, J.A.; Kumar, V. *Building Quality Management Systems: Selecting the Right Methods and Tools*; CRC Press: Boca Raton, FL, USA, 2013.
20. Ismyrlis, V. The contribution of quality tools and integration of quality management systems to the organization. *TQM J.* **2017**, *29*, 677–689. [CrossRef]
21. Wang, D.H.M.; Huynh, Q. The relationships among quality management system, knowledge management and organizational performance: An application of the Heckman two-step method. *J. Appl. Econ. Bus. Res.* **2014**, *4*, 235–245.
22. *ISO 9000*; Quality Management Systems Requirements. ISO: Geneva, Switzerland, 2015.
23. Soković, M.; Jovanović, J.; Krivokapić, Z.; Vujović, A. Basic quality tools in continuous improvement process. *J. Mech. Eng.* **2009**, *55*, 1–9.
24. Ilies, L.; Osoian, C.; Zaharie, M. Quality management system in higher education-employers approach. *Manag. Chall. Contemp. Soc. Proc.* **2010**, 75–79.
25. Arpaci, I. Antecedents and consequences of cloud computing adoption in education to achieve knowledge management. *Comput. Hum. Behav.* **2017**, *70*, 382–390. [CrossRef]
26. Lian, J.W. Critical factors for cloud-based e-invoice service adoption in Taiwan: An empirical study. *Int. J. Inf. Manag.* **2015**, *35*, 98–109. [CrossRef]
27. Nguyen, T.D.; Nguyen, T.M.; Pham, Q.T.; Misra, S. Acceptance and use of e-learning based on cloud computing: The role of consumer innovativeness. In *International Conference on Computational Science and Its Applications*; Springer: Cham, Switzerland, 2014; pp. 159–174. [CrossRef]
28. Harfoushi, O.; Akhorshaideh, A.H.; Aqqad, N.; Al Janini, M.; Obiedat, R. Factors affecting the intention of adopting cloud computing in Jordanian hospitals. *Commun. Netw.* **2016**, *8*, 88. [CrossRef]
29. Matar, N.; AlMalahmeh, T.; Sowan, B.; Al-Jaghoub, S.; Mater, W. A multi-group structural equation modeling for assessing behavioral intention of using mobile cloud computing-the case of Jordanian universities during the covid19 pandemic. *Int. Arab J. Inf. Technol.* **2022**, *19*, 203–221. [CrossRef]
30. Chang, V.; Walters, R.J.; Wills, G.B. Organisational sustainability modeling—An emerging service and analytics model for evaluating Cloud Computing adoption with two case studies. *Int. J. Inf. Manag.* **2016**, *36*, 167–179. [CrossRef]
31. Kumar, S.; Buyya, R. Green cloud computing and environmental sustainability. *Harnessing Green IT Princ. Pract.* **2012**, 315–339. [CrossRef]
32. Gill, S.S.; Buyya, R. A taxonomy and future directions for sustainable cloud computing: 360-degree view. *ACM Comput. Surv. (CSUR)* **2018**, *51*, 1–33. [CrossRef]
33. Gupta, P.; Seetharaman, A.; Raj, J.R. The usage and adoption of cloud computing by small and medium businesses. *Int. J. Inf. Manag.* **2013**, *33*, 861–874. [CrossRef]
34. Coccoli, M.; Guercio, A.; Maresca, P.; Stanganelli, L. Smarter universities: A vision for the fast-changing digital era. *J. Vis. Lang. Comput.* **2014**, *25*, 1003–1011. [CrossRef]
35. Sabi, H.M.; Uzoka, F.M.E.; Langmia, K.; Njeh, F.N. Conceptualizing a model for adoption of cloud computing in education. *Int. J. Inf. Manag.* **2016**, *36*, 183–191. [CrossRef]
36. Low, C.; Chen, Y.; Wu, M. Understanding the determinants of cloud computing adoption. *Ind. Manag. Data Syst.* **2011**, *111*, 1006–1023. [CrossRef]
37. Qian, L.; Luo, Z.; Du, Y.; Guo, L. Cloud computing: An overview. In *IEEE International Conference on Cloud Computing*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 626–631. [CrossRef]
38. Alotaibi, M.B. Exploring Users' Attitudes and Intentions toward the Adoption of Cloud Computing in Saudi Arabia: An Empirical Investigation. *JCS* **2014**, *10*, 2315–2329. [CrossRef]
39. Oliveira, T.; Thomas, M.; Espadanal, M. Assessing the determinants of cloud computing adoption: An analysis of the manufacturing and services sectors. *Inf. Manag.* **2014**, *51*, 497–510. [CrossRef]
40. Shon, T.; Cho, J.; Han, K.; Choi, H. Toward advanced mobile cloud computing for the internet of things: Current issues and future direction. *Mob. Netw. Appl.* **2014**, *19*, 404–413. [CrossRef]

41. Arpaci, I. Understanding and predicting students' intention to use mobile cloud storage services. *Comput. Hum. Behav.* **2016**, *58*, 150–157. [CrossRef]
42. Fishbein, M.; Ajzen, I. *Understanding Attitudes and Predicting Social Behavior*; Prentice-Hall: Hoboken, NJ, USA, 1980.
43. Fishbein, M.; Ajzen, I. *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*; Addison-Wesley: Reading, MA, USA, 1975.
44. Ajzen, I. The theory of planned behavior. *Organ. Behav. Hum. Decis. Process.* **1991**, *50*, 179–211. [CrossRef]
45. Venkatesh, V.; Morris, M.G.; Davis, G.B.; Davis, F.D. User acceptance of information technology: Toward a unified view. *MIS Q.* **2003**, *27*, 425–478. [CrossRef]
46. Venkatesh, V.; Goyal, S. Expectation disconfirmation and technology adoption: Polynomial modeling and response surface analysis. *MIS Q.* **2010**, *34*, 281–303. [CrossRef]
47. Venkatesh, V.; Thong, J.Y.; Xu, X. Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Q.* **2012**, *36*, 157–178. [CrossRef]
48. Thong, J.Y.; Hong, S.J.; Tam, K.Y. The effects of post-adoption beliefs on the expectation-confirmation model for information technology continuance. *Int. J. Hum. Comput. Stud.* **2006**, *64*, 799–810. [CrossRef]
49. Parasuraman, A.; Zeithaml, V.A.; Malhotra, A. ES-QUAL: A multiple-item scale for assessing electronic service quality. *J. Serv. Res.* **2005**, *7*, 213–233. [CrossRef]
50. Kim, Y.; Krishnan, R. On product-level uncertainty and online purchase behavior: An empirical analysis. *Manag. Sci.* **2015**, *61*, 2449–2467. [CrossRef]
51. Pahnila, S.; Mikko, S.; Viasong, Z. Integrated Habit into UTAUT: The Chinese eBay Case. *Pac.-Asia J. Assoc. Inf. Syst.* **2011**, *3*, 1–30.
52. Davis, F.D. A Technology Acceptance Model for Empirically Testing New End-User Information Systems Theory and Result. Ph.D. Dissertation, Massachusetts Institute of Technology, Sloan School of Management, Cambridge, MA, USA, 1986. Available online: https://dspace.mit.edu/bitstream/handle/1721.1/15192/14927137-MIT.pdf (accessed on 10 July 2022).
53. Lee, Y.; Kozar, K.A.; Larsen, K.R. The technology acceptance model: Past, present, and future. *Commun. Assoc. Inf. Syst.* **2003**, *12*, 50. [CrossRef]
54. Yaseen, S.G.; Dajani, D.; Hasan, Y. The impact of intellectual capital on the competitive advantage: Applied study in Jordanian telecommunication companies. *Comput. Hum. Behav.* **2016**, *62*, 168–175. [CrossRef]
55. Hew, T.S.; Kadir, S.L.S.A. Behavioral intention in cloud-based VLE: An extension to Channel Expansion Theory. *Comput. Hum. Behav.* **2016**, *64*, 9–20. [CrossRef]
56. Williams, M.D.; Rana, N.P.; Dwivedi, Y.K. The unified theory of acceptance and use of technology (UTAUT): A literature review. *J. Enterp. Inf. Manag.* **2015**, *28*, 443–488. [CrossRef]
57. Dajani, D.; Hegleh, A.S.A. Behavior intention of animation usage among university students. *Heliyon* **2019**, *5*, e02536. [CrossRef]
58. Abushakra, A.; Nikbin, D. Extending the UTAUT2 model to understand the entrepreneur acceptance and adoption internet of things (IoT). In Proceedings of the International Conference on Knowledge Management in Organizations (KMO 2019), Zamora, Spain, 15–18 July 2019; Springer: Cham, Switzerland, 2019; pp. 339–347.
59. El-Masri, M.; Tarhini, A. Factors affecting the adoption of e-learning systems in Qatar and USA: Extending the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2). *Educ. Technol. Res. Dev.* **2017**, *65*, 743–763. [CrossRef]
60. Alshirah, M.; Lutfi, A.; Alshirah, A.; Saad, M.; Ibrahim, N.M.E.S.; Mohammed, F. Influences of the environmental factors on the intention to adopt cloud-based accounting information system among SMEs in Jordan. *Accounting* **2021**, *7*, 645–654. [CrossRef]
61. Olson, J.M.; Zanna, M.P. Attitudes and attitudes change. *Annu. Rev. Psychol.* **1993**, *44*, 117. [CrossRef]
62. Wicker, A.W. Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. *J. Soc.* **1969**, *25*, 41–78. [CrossRef]
63. Hale, J.L.; Householder, B.J.; Greene, K.L. The theory of reasoned action. In *The Sage Handbook of Persuasion: Developments in Theory and Practice*, 2nd ed.; Dillard, J.P., Shen, L., Eds.; Sage Publications: Thousand Oaks, CA, USA, 2002; pp. 259–286. [CrossRef]
64. Faham, E.; Asghari, H. Determinants of behavioral intention to use e-textbooks: A study in Iran's agricultural sector. *Comput. Electron. Agric.* **2019**, *165*, 104935. [CrossRef]
65. Sparks, P.; Shepherd, R.; Wieringa, N.; Zimmermanns, N. Perceived behavioral control, unrealistic optimism and dietary change: An exploratory study. *Appetite* **1995**, *24*, 243–255. [CrossRef]
66. Banerjee, D.; Cronan, T.P.; Jones, T.W. Modeling it ethics: A study in situational ethics. *MIS Q.* **1998**, *22*, 31–60. [CrossRef]
67. Leonard, L.N.K.; Cronan, T.P.; Kreie, J. What influences ethical behavior Intentions planned behavior, reasoned action, perceived importance or individual characteristics? *Inf. Manag.* **2004**, *42*, 143–158. [CrossRef]
68. Chan, M.; Woon, I.; Kankanhalli, A. Perceptions of information security in the workplace: Linking information security climate to compliant behavior. *J. Inf. Priv. Secur.* **2005**, *1*, 18–41. [CrossRef]
69. Cox, J. Information systems user security: A structured model of the knowing-doing gap. *Comput. Hum. Behav.* **2012**, *28*, 1849–1858. [CrossRef]
70. Johnston, A.C.; Warkentin, M. Fear appeals and information security behaviors: An empirical study. *MIS Q.* **2010**, *34*, 549–566. [CrossRef]
71. Knapp, K.J.; Marshall, T.E.; Rainer, R.K.; Ford, F.N. Information security: Management's effect on culture and policy. *Inf. Manag. Comput. Secur.* **2006**, *14*, 24–36. [CrossRef]

72. Armitage, C.J.; Conner, M. Efficacy of the theory of planned behavior: Ameta-analytic review. *Br. J. Soc. Psychol.* **2001**, *40*, 471–499. [CrossRef] [PubMed]
73. Greaves, M.; Zibarras, L.D.; Stride, C. Using the theory of planned behavior to explore environmental behavioral intentions in the workplace. *J. Environ. Psychol.* **2013**, *34*, 109–120. [CrossRef]
74. Kidwell, B.; Jewell, R.D. An examination of perceived behavioral control: Internal and external influences on intention. *Psychol. Mark.* **2003**, *20*, 625–642. [CrossRef]
75. Lam, T.; Hsu, C.H. Predicting behavioral intention of choosing a travel destination. *Tour. Manag.* **2006**, *27*, 589–599. [CrossRef]
76. Alalwan, A.A.; Dwivedi, Y.K.; Rana, N.P. Factors influencing adoption of mobile banking by Jordanian bank customers: Extending UTAUT2 with trust. *Int. J. Inf. Manag.* **2017**, *37*, 99–110. [CrossRef]
77. Shiau, W.L.; Chau, P.Y. Understanding behavioral intention to use a cloud computing classroom: A multiple model comparison approach. *Inf. Manag.* **2016**, *53*, 355–365. [CrossRef]
78. Mahardika, H.; Thomas, D.; Ewing, M.T.; Japutra, A. Comparing the temporal stability of behavioral expectation and behavioral intention in the prediction of consumers' pro-environmental behavior. *J. Retail. Consum. Serv.* **2020**, *54*, 101943. [CrossRef]
79. Cheng, T.E.; Lam, D.Y.; Yeung, A.C. Adoption of internet banking: An empirical study in Hong Kong. *Decis. Support Syst.* **2006**, *42*, 1558–1572. [CrossRef]
80. Madden, T.J.; Ellen, P.S.; Ajzen, I. A comparison of the theory of planned behavior and the theory of reasoned action. *Personal. Soc. Psychol. Bull.* **1992**, *18*, 3–9. [CrossRef]
81. Wu, L.; Chen, J.L. An extension of trust and TAM model with TPB in the initial adoption of online tax: An empirical study. *Int. J. Hum. Comput. Stud.* **2005**, *62*, 784–808. [CrossRef]
82. Ringle, C.M.; Sarstedt, M. Gain more insight from your PLS-SEM results: The importance-performance map analysis. *Ind. Manag. Data Syst.* **2016**, *116*, 1865–1886. [CrossRef]
83. Hair, J.F., Jr.; Hult, G.T.M.; Ringle, C.M.; Sarstedt, M. *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*; Sage Publications: Thousand Oaks, CA, USA, 2021.
84. Michalos, A.C. (Ed.) . *Encyclopedia of Quality of Life and Well-Being Research*; Springer: Dordrecht, The Netherlands, 2014.
85. Cronbach, L.J. Coefficient alpha and the internal structure of tests. *Psychometrika* **1951**, *16*, 297–334. [CrossRef]
86. Netemeyer, R.G.; Bearden, W.O.; Sharma, S. *Scaling Procedures: Issues and Applications*; Sage Publications: Thousand Oaks, CA, USA, 2003.
87. Fornell, C.; Larcker, D. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *J. Mark. Res.* **1981**, *18*, 39–50. [CrossRef]
88. Hair, J.F.; Ringle, C.M.; Sarstedt, M. PLS-SEM: Indeed, a silver bullet. *J. Mark. Theory Pract.* **2011**, *19*, 139–152. [CrossRef]
89. Henseler, J.; Ringle, C.M.; Sarstedt, M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J. Acad. Mark. Sci.* **2015**, *43*, 115–135. [CrossRef]
90. Venkatesh, V.; Thong, J.Y.; Xu, X. Unified theory of acceptance and use of technology: A synthesis and the road ahead. *J. Assoc. Inf. Syst.* **2016**, *17*, 328–376. [CrossRef]
91. Sharif, A.; Afshan, S.; Qureshi, M.A. Acceptance of learning management system in university students: An integrating framework of modified UTAUT2 and TTF theories. *Int. J. Technol. Enhanc. Learn.* **2019**, *11*, 201–229. [CrossRef]
92. Yen, Y.; Wang, Z.; Shi, Y.; Xu, F.; Soeung, B.; Sohail, M.T.; Juma, S.A. The predictors of the behavioral intention to the use of urban green spaces: The perspectives of young residents in Phnom Penh, Cambodia. *Habitat Int.* **2017**, *64*, 98–108. [CrossRef]
93. Jafarkarimi, H.; Saadatdoost, R.; Sim, A.T.H.; Hee, J.M. Behavioral intention in social networking sites ethical dilemmas: An extended model based on theory of planned behavior. *Comput. Hum. Behav.* **2016**, *62*, 545–561. [CrossRef]
94. Gleim, M.R.; Smith, J.S.; Andrews, D.; Cronin, J.J., Jr. Against the green: A multi-method examination of the barriers to green consumption. *J. Retail.* **2013**, *89*, 44–61. [CrossRef]
95. Xu, X.; Wang, S.; Yu, Y. Consumer's intention to purchase green furniture: Do health consciousness and environmental awareness matter? *Sci. Total Environ.* **2020**, *704*, 135275. [CrossRef] [PubMed]
96. Sheppard, B.H.; Jon, H.; Warshaw, P.R. The theory of reasoned action: A meta-analysis of past research with recommendations for modifications and future research. *J. Consum. Res.* **1988**, *15*, 325–343. [CrossRef]
97. Davis, F.D.; Bagozzi, R.P.; Warshaw, P.R. User acceptance of computer technology: A comparison of two theoretical models. *Manag. Sci.* **1989**, *35*, 982–1003. [CrossRef]
98. Mathieson, K. Predicting user intentions: Comparing the technology acceptance model with the theory of planned behavior. *Inf. Syst. Res.* **1991**, *2*, 173–191. [CrossRef]

sustainability

MDPI

Article

Understanding Students' Perception of Sustainability: Educational NLP in the Analysis of Free Answers

Hiroko Yamano [1,*], John Jongho Park [2], Nathan Hyungsok Choe [3] and Ichiro Sakata [1,4]

1 Institute for Future Initiatives, The University of Tokyo, Tokyo 113-0033, Japan
2 School of Engineering Design, Technology, and Professional Programs, The Pennsylvania State University, University Park, PA 16802, USA
3 Department of Biomedical Engineering, The George Washington University, Washington, DC 20052, USA
4 Graduate School of Engineering, The University of Tokyo, Tokyo 113-8656, Japan
* Correspondence: yamano@ifi.u-tokyo.ac.jp; Tel.: +81-3-5841-8522

Abstract: This study explored undergraduate students' conceptions of sustainable development by asking about their definition of a sustainable world, current issues of sustainable development, and the necessary mindset and skillsets to build a sustainable world. We derived data from 107 participants' open-ended answers that we collected through an online survey at the beginning and the end of the sustainability class. Text mining with Natural Language Processing (NLP), principal component analysis (PCA), and co-occurrence network analysis were conducted to understand the changes in students' conception of sustainable development. In addition, we also conducted the Linguistic Inquiry and Word Count (LIWC) dictionary to investigate the psychometric properties of students' awareness and understanding related to sustainable development. This advanced analysis technique provided a rich understanding of university students' perceptions of sustainable development compared to what the UN initially defined as sustainable development goals (SDGs). The results showed imperative insights into the benefits of sustainability experiences and knowledge that generate motivation to develop students' competencies as change agents.

Keywords: natural language processing (NLP); sustainability education; sustainable development goals (SDGs); higher education; co-occurrence network; linguistic analysis

Citation: Yamano, H.; Park, J.J.; Choe, N.H.; Sakata, I. Understanding Students' Perception of Sustainability: Educational NLP in the Analysis of Free Answers. *Sustainability* **2022**, *14*, 13970. https://doi.org/10.3390/su142113970

Academic Editor: Changwoo Ahn

Received: 1 August 2022
Accepted: 3 October 2022
Published: 27 October 2022

1. Introduction

1.1. Definitions of Sustainability

The definition of sustainability has been elaborated to explain sustainability related to the human environment, referring to a technique or resource utilization [1]. In contrast, the resource is not exhausted or devastated in perpetuity. This inclusive definition brought about ambiguity, and vast flexibility applied to a broad spectrum of interests from various disciplinary fields. Thus, the interpretation of sustainability can vary depending on the purpose in a given area, however, the universal definition of sustainability applies to all scopes, disciplines, and facets of human effort. For instance, Moore et al. conducted a systematic literature review to conceptualize the definition of sustainability and fundamental concepts including the following categories: (1) after a defined period of time, (2) a program, clinical intervention, implementation strategies that continue to be delivered and (3) individual behavior change (i.e., clinician, patient) is maintained; (4) the program and individual behavior change may evolve or adapt while (5) continuing to produce benefits for individuals/systems [2]. This categorization across definitions reflects that there will be a desired sustained behavior maintained after a specified time an intervention/program is implemented.

Despite these varied definitions of sustainability, the clear focus is to develop solutions for sustainability issues, including growing concerns regarding pollution, social inequities, etc. Although these issues may seem daunting, sustainability goals have been developed to provide solutions to those crucial issues [3]. However, even though there has been a continuous effort to operationalize and conceptualize the concept of sustainability, the word "sustainability" remains ambiguous and multidimensional concept that is inherently difficult to incorporate into education [4].

Johnston et al. made an early attempt to operationalize the definition of sustainability while there were more than three hundred alternative and variously modified definitions of sustainability and sustainable development that exist broadly within the domain of environmental management [5]. If, however, true sustainability in human interactions with the biosphere is to be realized, a far more vital and more empirical interpretation of the original intent is urgently required. To be effective, such an interpretation must encompass and guide developments in political instruments and public policy as well as corporate decision-making and should focus increasingly on addressing the root causes of significant threats to sustainability rather than just their consequences.

Bilen et al. suggested the purpose of sustainability as (1) identifying a stable but inherently unjust equilibrium that causes the exclusion, marginalization, or suffering of a segment of humanity; (2) finding an opportunity in this unjust equilibrium, developing a social value proposition, and bringing to bear inspiration, thereby challenging the stable state's hegemony; and (3) forging a new, stable equilibrium that releases trapped potential or alleviates the suffering of the targeted group, and through imitation and the creation of a stable ecosystem ensuring a better future [6].

1.2. Education for Sustainable Development

The old paradigm of education for sustainable development that began in the 1960s and 70s focused primarily on ecology and the natural world. Since this was not a holistic approach considering other dimensions of care and motivation or resilience, the topics may not have resonated with students as they may not have seen its immediate relevance. More recently, the emerging academic field of sustainability focuses on introducing complex anthropogenic issues including climate change, desertification, poverty, pandemics, and war [7,8]. To address and resolve these sustainability challenges, different academic fields integrate and link existing collective knowledge of problems and solutions to academic settings so that critical sustainability issues will be solved [9,10].

The introduction of education for sustainable development was at the UN Rio Summit in 1992 [11]. During this summit, instructors and curriculum development were identified as significant areas to improve education for sustainable development. After 10 years, the UN General Assembly initiated the Decade of Education for Sustainable Development (DESD) which focused on expediting implementations of the UN sustainable development goals by developing and disseminating education for sustainable development [12].

More recently, following the UN 2030 Agenda for Sustainable Development guideline, the Global Action Programme (GAP) on education for sustainable development was introduced and implemented in 2015. This program explains that sustainability education should be at the center of the plan for sustainable development, and accelerate the transformation of sustainability education in higher education through research and improved teaching and learning environments of sustainability, mainly in developing countries [13].

Further, in 2019, UNESCO adopted the new global framework for education for sustainable development that focuses on integrating education to sustainable development into policies, learning environments, capacity building of educators, empowerment and mobilization, and local level action. The UN General Assembly noted that implementing education for sustainable development will be the critical aspect of accomplishing the UN sustainable development goals [14].

Therefore, to address and resolve sustainability challenges, different academic fields integrate and link existing collective knowledge of problems and solutions to academic settings so that critical sustainability issues, many scholars and educators began to investigate and develop education for sustainable development with pedagogical techniques that promote system thinking, service learning, participatory learning, and project-based education [8–10].

Considering complex sustainability issues, Sterling et al. introduced a holistic approach to sustainability education, seeing it as a process of maintaining a healthy system that draws on qualities such as creativity, self-reliance, self-realization, wholeness, and resilience [15]. Based on Sterling's holistic approach, Fisher et al. introduced three imperative dimensions including cultivating an ethic of care, fostering resilience and regeneration, and advocating for life's flourishing [16]. These concepts have been widely introduced for comprehensive sustainability education. They also emphasized the concept of cultivating an ethic of care that expresses the concept of "Care for the community of life with understanding, compassion, and love." Fisher and colleagues claimed that this aspect is immensely imperative for education for sustainable development as if there is no genuine care for self, others, nature and the earth, then there will not be any intrinsic motivation to maintain them.

1.3. Instructional Strategies for Sustainable Development Education

As for the instructional strategies of education for sustainable development, Remington and colleagues introduced the use of real-world problems in sustainability education [17]. This provides a vital context and relevance as well as allows students to solve real world problems and engage in a project in teams. In addition, project-based assignments that support students to work in teams to solve problems have become more common in recent times and have become more common in higher education [18,19]. Rulifson and Bielefeldt explored students' conceptualizations of sustainability by participating in sustainability projects [20]. They found that varied project experiences including social-learning shaped students' views of sustainability. These projects converged to develop students' understanding of socially responsible works to include safety, ethics, and various sustainability issues.

Smith et al. claimed that project-based learning empowers socio-technical thinking as an imperative element for engineers who will address sustainability issues [21]. They proposed that these sustainability-related projects should include integration of design thinking and interdisciplinary problem-solving into individuals' knowledge of global, societal, economic, and environmental contexts. Amadei and Wallace also mentioned that it is critical for academic programs to train students to develop their socio-technical perspectives so that they will be able to address global sustainability issues [22].

2. Research Rationale

The concept of sustainable development was first introduced by the United Nations in 1992 and expedited by various nations and organizations to implement the UN sustainable development goals. Ever since, there have been continuous efforts to operationalize the concept of sustainability to develop the current education curriculum that incorporates existing knowledge of sustainable development and current issues. In order to properly integrate sustainable development goals into current education, it is necessary to understand students' conceptions of sustainability and current issues of sustainable development.

While limited in their ability to compass all variables and factors, the use of conceptual frameworks is vital in the training and development process for sustainability education. . According to previous operational definitions and categorizations of sustainability, Wiek et al. identified critical competencies for sustainability change agents including systems-thinking, anticipatory, normative, strategic, and interpersonal competencies to develop a framework for sustainability-related academic programs [23]. The authors proposed that education for sustainability should be based on those competencies. They highlight the concepts of sustainability should be different by period, characteristics of stakeholders, temporal phases, uncertainty and epistemic status, inertia and path dependency, consistency and plausibility of future developments, and risk and intergenerational equity. They addressed that higher education sustainability courses should be designed based on key concepts and encourage students to acquire competencies. Regardless of a continuous effort to conceptualize the concept of sustainability, the meaning remains complex and multidimensional that is inherently difficult to comprehend [4,24]. Aligned with Withycombe et al.'s comment on operationalizing the concept of sustainability, to properly integrate the sustainable development goals into current education, it is necessary to understand students' conceptions of sustainability and current issues of sustainable development.

2.1. Target of Research

This study was conducted in a sustainability education class at a research-intensive university in Australia from February to November 2021. This class is designed to introduce students to the essential professional skills and mindsets for building a sustainable world. In this course, students learn how to lead and sustain themselves, and their relationships with others in our volatile, uncertain, complex, ambiguous (VUCA) world. The course has an emphasis on self-development and leadership based on a highly sophisticated framework extending awareness and understanding through empathy and systems thinking that leads to practical problem solving and successful project work. The focus areas for projects are based on the United Nations Sustainable Development Goals (SDGs). The course consists of weekly lectures covering essential topics on leadership and sustainability, tutorials where students engage in various activities supported by tutors and mentors, and mini courses led by domain experts to reflect on the points that can be applied to the team projects. The characteristic of this course is its focus on combining technical and non-technical skills for long-term effectiveness designed for high-achieving scholarship students. The course focuses on individual critical and creative thinking, reflection and contemplation, communication and collective creative problem-solving in a team, from a deep sense of awareness and understanding. This course seeks a shift in mindset, a transformation to see sustainable development problems differently, to relate to oneself, others, and the world in a new way of considering the root causes of issues with a holistic whole system perspective [25].

2.2. Quantitative Analysis of Students' Free Answers

Course surveys are essential tools for evaluating and updating the class contents. The accessible answers of students, in particular, often contain valuable information and provide an essential basis for rebuilding current education. They are helpful in evaluating classes targeting diverse and multi-domain issues as seen in sustainability education. Previous researches that analyzed free answers to investigating the changes in students' perceptions and attitudes through the course have been primarily qualitative [26–28]. However, the summaries and extractions that human analysts make of these open-ended answers tend to rely on intuition and are not satisfactorily reliable [29]. Given the recent increase of massive open online courses during the pandemic, it has become easier to take surveys and receive feedback on a large scale, which requires precision and agility for the evaluation [30]. Therefore, applying computer-based quantitative analysis that can automatically mine useful information from open answers has become an important issue. In this study, we applied an integrated text mining approach with Natural Language Processing (NLP), principal component analysis (PCA), and co-occurrence network analysis to understand the changes in students' conception of sustainability from their free answers. We also used Linguistic Inquiry and Word Count (LIWC) dictionary to investigate the psychometric properties of students' awareness and understanding related to sustainability.

2.3. Research Purpose

The purpose of this study is to verify whether the course would achieve expected effect as a sustainability education. Based on the characteristics of the course that focuses on the transformation in students' conception, this study sought to analyze the changes in the basic psychometric properties of students' awareness and understanding related to the UN sustainable development goals. On two occasions before and after the class, participants were asked to answer their standpoints toward sustainability. Of the 32 questions executed in the survey, we focused on the three descriptive questions and their open-ended answers as follows, which form the core of the survey.

Q1. What is your definition of a sustainable world?
Q2. What kind of sustainability issues are critical and should be addressed in creating a sustainable world? Please explain your standpoint based on your experience.
Q3. What are the essential mindset and skills required to build a sustainable world? Please explain why you would need such a mindset and skills.

While there were limitations to surveying all questions and responses, the use of free answers is vital for examining the effectiveness of the sustainability course that addresses complex issues as we discussed above. Three broad research questions around relative differences in students' understanding and language use in their answers guided the analyses: (a) How different is students' answer distribution before and after the class? (b) How distinctive is the word structure in students' answers? (c) How different is students' language use across questions before and after the class? Based on these research interests, we formulated the following hypotheses regarding the main trends in the changes in students' responses to each question after the class.

H1. Students obtained diverse aspects regarding the definition of sustainability. (Q1)
H2. Students shared consolidated views on the critical issues of sustainability. (Q2)
H3. Student opinions of the essential mindset and skills for a sustainable world remained diverse without common understanding. (Q3)

To shed light on the results, we used text mining techniques to investigate undergraduate students' conceptions of sustainability and necessary mindsets and skills. Text mining is the process of exploring and analyzing text data that includes various methods such as NLP, PCA, and co-occurrence network analysis, explained in the next section. This advanced analysis technique provided a rich understanding of university students' perspectives on sustainability compared to the UN originally defined sustainable development goals.

3. Materials and Methods

3.1. Participants

Participants were scholarship students enrolled in a sustainability class at a research-intensive university in Australia who scored in the top 2% of Australian Tertiary Admission Rank scores. Students come from various disciplines including Engineering, Business, Science and Arts. A total of 107 out of 122 students completed the survey comprising 67 males (54.9%) and 55 females (45.1%). Most were in their first year (81, 66.4%), with 24 second-year (19.7%), 16 third-year (13.1%), and 1 fourth-year student. There were 33 Americans (27%), 23 Sub-Saharan Africans (18.9 %), 21 Asians (17.2%), 13 Oceanians (10.7%), 12 North African and Middle Eastern (9.8%). 14 Europeans (9), and others who participated in this study. Half of the students majored in Engineering (64, 52.4 %), 20 and 17 of them majored in Arts and Social Science (16.4%) and Architecture, Design & Planning (13.9%) as follows. When they were asked about their career intention in STEM, 84 students (68.9%) answered that they planned to pursue a STEM-related career. Data sources included responses to open-ended questions delivered via a Qualtrics survey about building a sustainable world. The student characteristics included demographic and academic information such as gender, ethnicity, major, and program year. Participation was voluntary, and students could decide to stop participating in the survey at anytime. Students did not have to answer any questions they did not want to answer. Their participation did not have any impact on their grades or engagement in the course. Their confidentiality was maintained to the degree permitted by the technology used and we did not capture identifiable information.

3.2. Data Analysis Methods

In this study, we conducted a comparison of students' free answers in before and after surveys using the techniques of NLP to prepare the free answers for analysis, as part of the pre-processing stage. NLP has been used for the text analysis of free-style answers, such as feature extraction [31,32], text classification [33], and sentiment analysis [34]. NLP is a set of methods to enable computers to understand the meaning of natural language, such as documents and sentences used by humans. It includes morphological analysis, syntactic analysis, semantic analysis, and contextual analysis. Morphological analysis is a method that divides sentences into morphemes, the minor linguistic units having meaning. It assigns information to morphemes such as parts of speech (POS tagging) using a machine-readable dictionary and a corpus of documents that stores the use of a language. In the pre-processing of documents, the stopwords unrelated to the topic are often removed beforehand. Stemming is a method that converts derivatives to the same feature. As many stemming methods use simple lexical processing, the converted result can have a different meaning. Therefore, the process of restoring words in their basic form is often used instead of stemming, which is called lemmatization. After the pre-processing, vector representation is often used to represent natural language in a mathematically tractable form that a computer can understand. Each element of a vector represents the features of a document or sentence. Each dimension of the vector is associated with a word, and its value is usually the frequency of words in the document. This kind of vector representation method of documents is called bag-of-words, as information about sentence structure and word order is lost. However, in many cases, the frequency of the word-based approach is sufficient to classify sentences according to their topics. In this study, since the scale of the data is relatively small, we used morphological analysis and vector representation based on the frequency of words.

To analyze the open-ended responses to the three questions, we executed basic pre-processing using Natural Language Toolkit (NLTK) [35] as the following: (1) Replace blank answers with empty space, (2) Lowercase all text, (3) Remove all blocks of digits, (4) Remove all punctuation and replace it with a space, (5) Remove all diacritics and accents, (6) Remove ubiquitous words using NLTK's English stopwords of 179 words, (7) Remove any extra whitespace, newline, tabs and any form of space. Next, to remove affixes and map various forms of a word to the canonical word in a dictionary (known as lemmatization), and to obtain the grammatical parts of speech tagging, we used a language model of en_core_web_sm provided by spaCy [36], the open-source library for advanced Natural Language Processing. en_core_web_sm is a trained pipeline for English that incorporates lexical databases from OntoNotes 5 [37] and WordNet 3.0 [38]. In the following analysis, we only used pre-processed words that are tagged as adjectives, verbs, nouns, proper nouns, or adverbs and excluded words tagged otherwise from the analysis data.

3.2.1. Principal Component Analysis (PCA)

To extract principal information of the answers to each question, we performed PCA [39] and students' answers were transformed into vector representations with the following procedures. We used each pre-processed word in a response as a "term", and each response in the target question as a "document". We calculated the values of terms in each document using TF–IDF [40], which is the product of the term frequency (TF) and the inverse document frequency (IDF). This scoring method is empirically useful for a wide range of datasets and has been proven to represent "the amount of information of a term weighted by its occurrence probability" [41]. TF–IDF is defined as the following equation,

$$\text{tfidf}(t, d) = \text{tf}(t, d) \cdot \text{idf}(t), \quad (1)$$
$$\text{tf}(t, d) = \log(1 + f_{t,d}), \quad (2)$$
$$\text{idf}(t) = \log \frac{N}{df_t}, \quad (3)$$

where t and d represent the term and the document, respectively, $\text{tf}(t, d)$ is the log value of the TF, and $\text{idf}(t)$ is the inverse ratio of the number of documents that include the term t. N represents the total number of documents. When calculating TF–IDF we ignored terms that have a document frequency lower than the threshold of 1.

By concatenating the values of terms in each response to each question, we obtained the vector representation of each question, which is the multi-dimensional array that corresponds to the TF–IDF values of the terms in the documents. To investigate the difference in the vector representation of the questions, we applied PCA. PCA is the dimensionality-reduction method that transforms variables into lower-dimensional data while preserving as much of the data's variation as possible, by decomposing data into two orthogonal matrices with maximum variance and singular values known as an eigenvalue. In this study, we reduced data into two principal components to visualize the vector representation of the questions into X,Y two-dimensional coordinates.

3.2.2. Co-Occurrence Network Analysis

A document is a set of words, and the meaning it contains can be regarded as being generated by the co-occurrence of words. Since the co-occurrence of words differs according to the documents, the semantic aspects of the topics they represent also change. In other words, topic context can be modeled using the co-occurrence relationships between the words in the topic [42]. A co-occurrence network is one of the primary methods of text mining, which visualizes such relationships between words in a document as a network, based on the similarity of word occurrence patterns in the sentences. The nodes in the network represent words, and the width of edges between nodes represents the strength of word relationships, i.e., the number of their co-occurrences. Many researchers have used a co-occurrence network of free answers, such as co-occurring emotions related to students dropping out during the Massive Open Online Course (MOOC) [43], self-reported labels of gender and sexual identities [44]. Co-word maps based on word occurrences are more effective than topic modeling for designating semantic components, especially for small and medium-sized data sets [45].

To capture the difference in the students' responses before and after the class, we counted the co-occurrences of words in every response, and the number of occurrences of all combinations of the pre-processed word pair except for pairs of the same word in response. Then we summed up the number of co-occurrences of pairs in the answers to each question, which corresponds to the weight of the edges in the co-occurrence network. For visualization, we selected the top 200 co-occurrence pairs and maximum connected components in the network to avoid complexity. The network was drawn using the Fruchterman-Reingold algorithm that simulates a force-directed representation of the network treating edges as springs holding nodes close while treating nodes as repelling objects. Simulation continues until the positions are close to an equilibrium. In this algorithm, nodes connected by edges are placed close together (attractive force), while nodes not connected by edges are placed farther apart (repulsive force). Therefore, in the case of a co-occurrence network, words with a high co-occurrence, which are the words students frequently use together in their responses, are placed closer and vice versa.

3.2.3. Linguistic Analysis with LIWC Dictionary

To investigate the psychometric properties of students' understanding related to sustainability, we investigated the linguistic features of their free answers, which employed the LIWC2015 program (LIWC–Linguistic Inquiry and Word Count). This application was created to serve as an efficient and effective method for studying the various emotional, cognitive, and structural components present in individual's verbal samples. It has been used to analyse students' self-descriptions associated with their properties such as depression and depression-vulnerability [46], course performance [47], and personality differences [48]. The LIWC was firstly developed as part of an exploratory study of language and disclosure [49] and has been constantly updated over time [50–52]. LIWC2015 software relies on an internal default dictionary defining which words should be counted in the target text. In this study, LIWC2015 English dictionary data was provided by the authors under the particular license they offered for more efficient analysis. LIWC2015 dictionary is composed of almost 6400 English words and 73 categories including linguistic dimensions, psychological constructs, and personal concerns.

Among the psychological categories in LIWC2015, we focused on four categories of Affective, Cognitive, Drives, and Relativity that showed large differences in the number of word occurrences before and after class. The four categories contain sub-categories as follows: Affective (Positive emotion, Negative emotion, Anxiety, Anger, Sadness), Cognitive (Insight, Causation, Discrepancy, Tentative, Certainty, Differentiation), Drives (Affiliation, Achievement, Power, Reward, Risk), and Relativity (Motion, Space, Time). Therefore, we used 4 categories and 19 sub-categories in total, to extract the changes of students' perceptions before and after the sustainability class.

4. Results

Data summary for students' responses is provided in Table 1. Since each question was not mandatory, the number of valid answers differed between questions before and after the sustainability class. The average word per answer indicates the average number of pre-processed words (adjectives, verbs, nouns, proper nouns, or adverbs) in each questionnaire. Unique pairs are the number of combinations of pre-processed words in the questionnaires' responses, which were made for the co-occurrence network. Normalized unique pairs were calculated by dividing individual pairs by valid answers and average words per answer. There were fewer valid responses and fewer pre-processed words and unique pairs in the responses after the class. Q2 had the highest valid responses, average number of words, and unique pairs among the three questions. Individual pairs in Q2 and the rate of their decrease after class was higher than those in Q1 and Q3. However, the values of the normalized unique pairs were not significantly different between questions.

Table 1. Data Summary.

	Valid Answers		Average Words per Answer		Unique Pairs		Normalized Unique Pairs	
Question	**Before**	**After**	**Before**	**After**	**Before**	**After**	**Before**	**After**
Q1	167	115	13.98	13.78	15,864	11,385	6.795	7.184
Q2	178	133	29.01	20.45	83,399	36,664	16.151	13.480
Q3	169	124	23.20	19.31	43,435	26,142	11.078	10.918
Average	171	124	22.06	17.85	47,566	24,730	12.585	11.173

Figure 1 plotting points are calculated by PCA analysis for each response to questions Q1, Q2, and Q3, respectively. The blue dots show the distribution of participants' responses at the beginning of the semester. The orange dots provide the distribution of participants' answers at the end of the semester. In Q3, there is not much significant change in the structure of the latent space between the blue and orange dots, indicating that participants tended to provide dispersed answers without any common concepts. However, in Q1 and Q2, the distributions of the dots are particularly clearly differentiated, indicating that their answers tended to be semantically different after the semester. Regarding the content of the questions, the answers to Q2 about "most critical sustainability issues" tended to converge. In contrast, the answers to Q1 about the "definition of sustainability" grew to become diverse after the class. These results show that PCA analysis seemed useful for distinguishing the significant changes in the respondents' answers before and after the course, especially concerning the urgency and definition of sustainability. However, it may not help detect minor or diversified changes in mindsets and skills needed for a sustainable world.

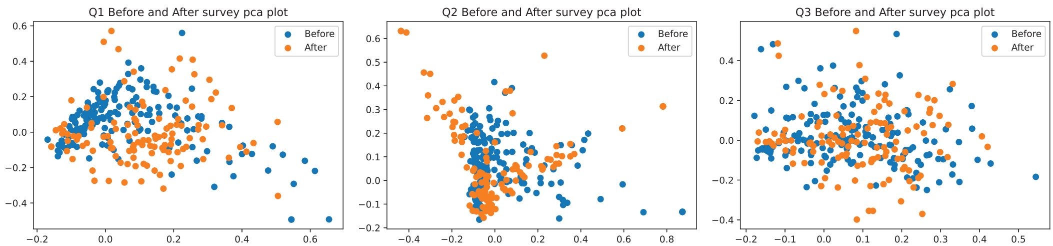


Figure 1. (**Q1.**) The distribution of responses on "What is your definition of a sustainable world? Please explain your stand point based on your experience." in the latent semantic space, both for semester beginning (blue) and semester-end (orange). (**Q2.**) The distribution of responses on "What kind of sustainability issues are most critical and should be addressed in creating a sustainable world? Please explain your stand point based on your experience." in the latent semantic space, both for semester beginning (blue) and semester-end (orange). (**Q3.**) The distribution of responses on "What are the essential mindset and skills that you would need to build a sustainable world? Please explain why you would need such a mindset and skills. ? Please explain your stand point based on your experience." in the latent semantic space, both for semester beginning (blue) and semester end (orange).

Figure 2 represents the co-occurrence network of the words in the responses to each question. The structure of the co-occurrence networks differed between the six networks of the questions surveyed before and after the class. Q1 network had one center, i.e., almost every word was connected to the centered two words of "sustainable" and "world". On the other hand, Q2 and Q3 networks had broader centers and the centered words were dispersed after the class. Especially in Q2, the center was divided into two centers with the appearance of the new hub of the word "poverty".

Of the words used in the responses, the coverage rate of the terms registered in the LIWC2015 dictionary was about 60%. Table 2 shows the number of pre-processed words in each question, registered words, and registration rate in the LIWC2015 dictionary, both in all categories (73 categories) and psychological categories (46 categories).

Table 2. A number of words and their registrations in LIWC2015 dictionary.

		73 Categories		46 Categories	
Group	**Words**	**Registered Words**	**Registration Rate**	**Registered Words**	**Registration Rate**
Q1 before	2335	1384	0.593	1311	0.561
Q2 before	5164	2936	0.569	2728	0.528
Q3 before	3921	2366	0.603	2252	0.574
Q1 after	1585	1009	0.637	969	0.611
Q2 after	2720	1590	0.585	1482	0.545
Q3 after	2394	1511	0.631	1431	0.598

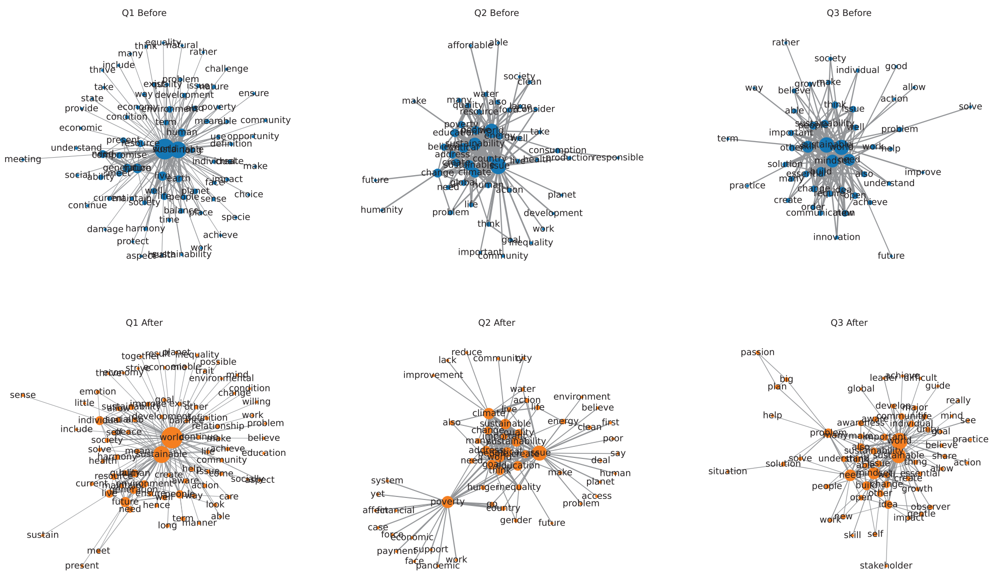


Figure 2. The co-occurrence network of the words in the responses to each question.

Figure 3 represents the Pearson correlation among the questions calculated by the distribution of word occurrences in psychological categories. We used 46 psychological categories, including sub-categories, in the LIWC2015 dictionary and compared the correlation of word distribution among the categories before and after the class. In total, highly correlated pairs were the group of the answers to the same questions before and after class, e.g., Q1_before and Q1_after.

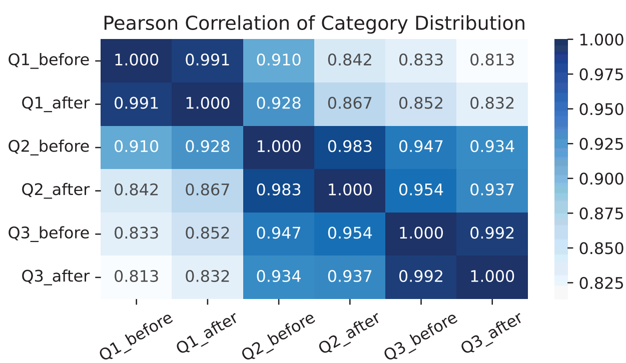


Figure 3. Pearson correlation of psychological category distribution.

The number of occurrences of the word in the targeted four psychological categories (Drives, Relativity, Affective, Cognitive), including 17 sub-categories in each question, is shown in Table 3. In LIWC2015, each word was assigned to approximately three categories on average, so the numbers were higher than the actual number of word occurrences. The values ranged from about 300 to 1650 with no significant bias in the distribution across categories.

Table 3. A number of word occurrences in psychological categories.

Group	Drive	Relativity	Affective	Cognitive
Q1 before	508	1012	331	504
Q1 after	426	731	303	358
Q2 before	1475	1639	982	1445
Q2 after	964	760	542	823
Q3 before	1446	1212	767	1777
Q3 after	872	739	523	1214

To compare the difference among categories, we calculated standard scores for each question before and after the class as the rate of the word occurrences in each of four categories with sub-categories, divided by their total number of occurrences in 43 psychological categories. Table 4 shows the representative word examples and their number of occurrences in the categories with significant differences in the standard scores before and after the class. The differences were calculated by the standard score before class minus those after class. If the difference was positive, before class word samples were displayed; if negative, after class samples were shown in Table 4. Among the categories, main categories were indicated with the asterisk "*". Achievement and Power were the sub-categories of Drive, and Work was the sub-category of Personal concerns. Therefore, we focused on the four categories of Drives, Relativity, Affective and Cognitive processes, and their sub-categories with high differences.

Table 4. Representative words in categories with significant difference of occurrences.

	Category	Difference	Examples
Q1	Affective *	−0.434	(well, 15), (peace, 9), (create, 7), (problem, 6), (support, 5)
Q1	Present focus	0.305	(live, 58), (need, 48), (meet, 21), (present, 15), (take, 10),
Q1	Discrepancy	0.296	(need, 48), (would, 12), (problem, 5), (must, 4), (rather, 3)
Q1	Social *	−0.277	(people, 40), (human, 19), (individual, 15), (help, 9), (social, 9)
Q1	Drives *	−0.277	(able, 13), (help, 9), (social, 9), (allow, 9), (meet, 8)
Q2	Relativity *	**0.814**	(world, 121), (change, 71), (global, 31), (action, 27), (environment, 22)
Q2	Drives *	**−0.613**	(education, 61), (goal, 37), (important, 32), (poverty, 29), (create, 24)
Q2	Power	−0.432	(education, 61), (important, 32), (poverty, 29), (help, 7), (big, 6)
Q2	Work	−0.424	(education, 61), (goal, 37), (work, 12), (achieve, 11), (resource, 10)
Q2	Time	0.382	(time, 17), (due, 12), (term, 11), (current, 11), (still, 10)
Q3	Achievement	0.433	(skill, 90), (work, 28), (solution, 27), (achieve, 21), (able, 20)
Q3	Work	0.354	(skill, 90), (work, 28), (achieve, 21), (learn, 14), (goal, 14)
Q3	Cognitive *	−0.311	(need, 56), (think, 31), (problem, 29), (change, 25), (other, 24)
Q3	Insight	−0.292	(think, 31), (idea, 19), (solution, 18), (understand, 15), (solve, 14)
Q3	Drives *	0.240	(skill, 90), (important, 39), (work, 28), (problem, 28), (solution, 27)

Figure 4 shows a scatter plot of the standard scores with the (X and Y) axes corresponding to (Drives, Relativity) with blue dots and (Affective, Cognitive) with orange dots, respectively. The red arrows indicate the most significant change in the after-class values compared to the before-class values, which corresponds to (Drives, Relativity) in Q2. On the contrary, the position of (Affective, Cognitive) in Q2 was almost the same before and after the class.

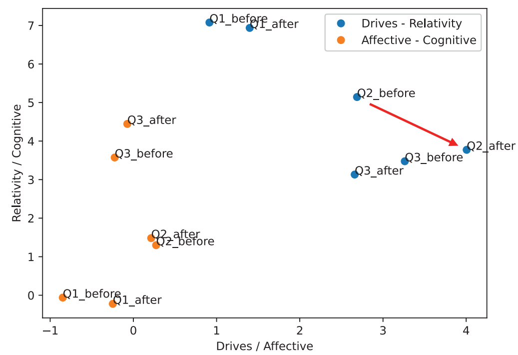


Figure 4. Changes in the psychology categories in the responses to each question before and after the class.

5. Discussion

This study aimed to determine the basic psychometric properties of students' understanding of sustainability as measured by their language use in their answers. This paper does not cover the full range of sustainability education. Instead, it focuses on the self-development and shift in mindset needed to address real-world sustainability concerns described in Section 2.1. Therefore, the primary purpose of this study was to evaluate the transformation in students' perception rather than memorizing knowledge about sustainability.

Due to the limited number of data and the variety of students' answers, we first focused on extracting major components of the data and their changes after the class. Then we confirmed the contents of the changes using network analysis and dictionary matching. Using network analysis in the text mining of free answers, we detected the unique perspective of the changes in students' responses after the course. Results show that students' understanding of critical issues has increased significantly, strengthening their motivation to build a sustainable world. According to the results, sustainability education is more than simply transferring knowledge to students. It is also driven by encouragement to acquire competencies for sustainability change agents, such as systems thinking, anticipatory, normative, strategic, and interpersonal competencies to cope with real-world problems [4].

5.1. Students' Understanding of Sustainability from Word Distribution

5.1.1. PCA Plotting

One of the goals of this study was to estimate fundamental changes in students' awareness and understanding of sustainability from the perspective of word distribution. Using TF-IDF values of words, we obtained the distributed representation in students' free answers and compared their differences using PCA plotting. Although PCA plotting is a naive method to visualize data points by summarizing word representation into two dimensions, it is essential to note that we found different trends among the answers. Since PCA extracts principal components with the most considerable variance in the data set, it reflects the main differences in the students' use of words. The distribution of the data points showed that students tended to give semantically different answers for Q1 and Q2 after class, with Q1 more dispersed and Q2 more aggregated. The Q1 is about the definition of a sustainable world; these dispersed Q1 PCA plots can be explained from class contents. This class provided UN 17 sustainable development goals to students. Thus, the students may expand their thoughts about definitions of sustainability related to the 17 goals to make a sustainable world. In addition, they gave continuously diverse answers over time in Q3, suggesting that students with different majors may provide different essential mindsets and skills to build a sustainable world. The results suggested that hypotheses H1, H2, and H3 were supported.

5.1.2. Co-Occurrence Network Analysis

Before assessing the hypotheses, we first observed the consistency between network structure and word use in students' answers. We applied co-occurrence network analysis by linking frequently occurring words in response, the top 200 frequent word pairs in this study. The typical star structure of the co-occurrence network of Q1 was consistent with the grammatical characteristic of English that requires a subject to complete a sentence. Since Q1 required the same subject to explain its definition, almost all answers contained "sustainable world" as a subject, resulting in a hub of two words in the network. On the other hand, Q2, and Q3 allowed for a wider variety of subjects and words in their responses, consistent with the network structure with a broader center. This network characteristic was aligned with the more prominent average words and unique pairs of Q2 and Q3, as shown in Table 1. From these observations, we assumed co-occurrence network analysis is functional for assessing the basic features of the answers formulated in the hypotheses. The emergence of the new hub "poverty" and its neighbor words' low number of links in Q2 after the class showed the uniqueness of the neighbors, suggesting that quite a few students tended to use more specific terms related to poverty after the course. These findings support our hypothesis H2, indicating that a substantial number of students shared the perception of poverty as a critical issue of sustainability. This result aligns with a few empirical studies on college students' perceptions of sustainability and poverty. For example, Watson et al. reported that engineering students participating in learning-cycle-based sustainability class were able to connect relations between sustainability and poverty well [53]. However, we could not find particular patterns that strongly explain hypotheses H1 and H3 from the analysis based on the overall structure of the network. For a more detailed assessment, we investigated the word distribution using linguistic analysis based on psychological categories as the following.

5.2. Psychological Transfer in Students' Understanding of Sustainability

5.2.1. Assessing the Stability

In studies analyzing time-series changes in spoken words, the stability of the term used was often a key issue concerning the reliability of the analysis. Stability has been investigated by the correlation or overlap of people's language use over time [54,55]. Based on the preceding studies related to stability, we investigated the Pearson correlation of the students' word use before and after the class to assess the reliability of the analysis (Figure 2). As a result, the groups of answers to the same question were highly correlated compared to the other questions, indicating stability in the tendency of students to give similar answers to the same questions. The stability among questions was also confirmed by the scatter plot of the standard scores in Figure 4, i.e., the positions of each question were relatively close before and after class compared to those of the other questions.

5.2.2. Assessing the Change

Among the relatively unchanged positions in the same questions, Q2, given Drive and Relativity, showed the most considerable shift before and after the class, contrary to Q2, given Affective and Cognitive process that showed the slightest shift (Figure 4). The LIWC2015 dictionary has five sub-categories within the Drive category: Affinity, Achievement, Power, Reward, and Risk, which represent the motivation factors that lead to behaviors. The new hub of "poverty" on the co-occurrence network in Figure 3 belongs to the sub-category of Power. Therefore, the hub "poverty" appears to emerge, accompanied by an increase in Drive-related words. Considering that the content of Q2 was "What are the essential issues for creating a sustainable world?" students' perceptions seemed to have deepened to produce more concrete actions rather than internal aspects such as emotion and cognition.

6. Conclusions

This study investigated students' perceptions of sustainability and the skills required to build a sustainable world from an educational perspective. The results showed imperative insights into the benefits of sustainability experiences and knowledge that generate motivation to develop students' competencies as change agents. In addition, this study's surveys were conducted at the beginning and the end of the sustainability class to explore changes in students' perceptions of building a sustainable world. The results indicated that the sustainability class impacts college students' perception of critical sustainability issues and these solutions. The word distribution PCA plots showed that class might assist students with understanding or identifying the core critical sustainability issues, given that the plots converged. Furthermore, a co-occurrence network analysis showed that students view poverty as an essential issue in their post-survey than in their pre-survey. In representative words in categories (Table 4), students used more drive words in the post-survey compared to the pre-survey and fewer relativity words in the post-survey compared to the pre-survey. Perhaps students have responded by providing solutions to build a sustainable world rather than raising sustainability issues. Lastly, the central phenomenon portrayed in Figure 2 shows that emerging poverty as a critical issue for sustainability tended to depend on motivations, as discussed above. These motivations were not limited to acquiring knowledge about sustainability. Thus, there were ways that the various interests of sustainability and envisioning a possible future self were connected to the sub-themes of sustainability. Based on our study, we recommend sustainability educators and practitioners include real-world-based projects that incorporate the United Nations' sustainable development goals and provide activities that address the fundamental issues with a holistic whole system perspective. This practice may help college students think comprehensively about sustainability issues and how to develop a sustainable world.

There are several caveats to our analysis. First, because the survey data used in this study were from the sustainability course designed for personal and leadership development from a holistic perspective and mutual understanding, it did not pursue domain-specific knowledge or techniques, as discussed in Section 5. Therefore, programs aimed at acquiring specific techniques that contribute to sustainability may not be applicable or produce different results for this study. Second, the visualizations of Figures 1 and 2 does not cover the entire data. Due to the limited number and variety of data in this study, it required more than 50 dimensions to meet the sufficient cumulative contribution rate above 0.6 for PCA validation. Therefore, the PCA plotting of 2-dimensional space does not represent the entire data distribution, and we used the visualization only to capture the major difference in the data. The characteristics of the overall structure of the network are also difficult to verify. Therefore, we visualized the network multiple times and confirmed its major characteristics did not change with the fluctuation caused by the visualization. PCA and co-occurrence networks only capture the principal differences among students' answers; however, the LIWC dictionary covers the majority of data, as shown in Table 2. Third, because we focused on the primary analysis based on word occurrences, the context of each student's answer was lost. In addition, because the data were not linked to individual names or IDs, the analysis result did not capture the individual student differences throughout the course. Applications of advanced NLP techniques, such as deriving domain models from free answers, are left for future research.

Author Contributions: Conceptualization, H.Y. and J.J.P.; methodology, H.Y.; software, H.Y.; formal analysis, H.Y.; investigation, J.J.P. and N.H.C.; resources, H.Y. and J.J.P.; data curation, H.Y.; writing—original draft preparation, H.Y. and J.J.P.; writing—review and editing, H.Y., J.J.P., N.H.C. and I.S.; visualization, H.Y.; project administration, N.H.C.; funding acquisition, H.Y. and I.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Acknowledgments: We thank Ehssan Sakhaee and Melissa Neighbour from the University of Sydney for data collection, and helpful discussion and revision of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Berkes, F.; Folke, C. Linking social and ecological systems for resilience and sustainability. *Link. Soc. Ecol. Syst. Manag. Pract. Soc. Mech. Build. Resil.* **1998**, *1*, 4.
2. Moore, J.E.; Mascarenhas, A.; Bain, J.; Straus, S.E. Developing a comprehensive definition of sustainability. *Implement. Sci.* **2017**, *12*, 110. [CrossRef]
3. Degirmenci, K.; Recker, J. Creating environmental sensemaking through green IS: An experimental study on eco-nudging paper printing behavior. In Proceedings of the 24th Americas Conference on Information Systems, Association for Information Systems (AIS), New Orleans, LA, USA, 16–18 August 2018; pp. 1–10.
4. Ngo, T.T.; Chase, B. Students' attitude toward sustainability and humanitarian engineering education using project-based and international field learning pedagogies. *Int. J. Sustain. High. Educ.* **2020**, *22*, 254–273. [CrossRef]
5. Johnston, P.; Everard, M.; Santillo, D.; Robèrt, K.H. Reclaiming the definition of sustainability. *Environ. Sci. Pollut. Res. Int.* **2007**, *14*, 60–66. [CrossRef] [PubMed]
6. Bilén, S.G.; Kisenwether, E.C.; Rzasa, S.E.; Wise, J.C. Developing and assessing students' entrepreneurial skills and mindset. *J. Eng. Educ.* **2005**, *94*, 233–243. [CrossRef]
7. Bäckstrand, K. Civic Science for Sustainability: Reframing the Role of Experts, Policy-Makers and Citizens in Environmental Governance. *Glob. Environ. Politics* **2003**, *3*, 24–41. [CrossRef]
8. Grunwald, A. Strategic knowledge for sustainable development: The need for reflexivity and learning at the interface between science and society. *Int. J. Foresight Innov. Policy* **2004**, *1*, 150. [CrossRef]
9. Bammer, G. Guiding Principles for Integration in Natural Resource Management (NRM) as a Contribution to Sustainability. *Australas. J. Environ. Manag.* **2005**, *12*, 5–7. [CrossRef]
10. Kerkhoff, L.v.; Lebel, L. Linking Knowledge and Action for Sustainable Development. *Annu. Rev. Environ. Resour.* **2006**, *31*, 445–477. [CrossRef]
11. Weiss, E.B. United Nations Conference on Environment and Development. *Int. Leg. Mater.* **1992**, *31*, 814–817. [CrossRef]
12. ESD Section, UNESCO. UN Decade of Education for Sustainable Development: The First Two Years. *J. Educ. Sustain. Dev.* **2007**, *1*, 117–126. [CrossRef]
13. Leicht, A.; Heiss, J.; Byun, W.J. *Issues and Trends in Education for Sustainable Development*; UNESCO Publishing: Paris, France, 2018; Volume 5.
14. Purnamasari, S.; Suhendi, F.A.F.; Zulfah, N.L.N. Implementasi Education for Sustainable Development (ESD) dalam pembelajaran IPA di Kabupaten Garut: Sebuah studi pendahuluan. *J. Kaji. Pendidik. IPA* **2022**, *2*, 7–12. [CrossRef]
15. Sterling, S.; Orr, D. *Sustainable Education: Re-Visioning Learning and Change*; Green Books for the Schumacher Society: Totnes, UK, 2001; Volume 6.
16. Fisher, J.R.; Wood, S.A.; Bradford, M.A.; Kelsey, T.R. Improving scientific impact: How to practice science that influences environmental policy and management. *Conserv. Sci. Pract.* **2020**, *2*, e210. [CrossRef]
17. Remington-Doucette, S.M.; Connell, K.Y.H.; Armstrong, C.M.; Musgrove, S.L. Assessing sustainability education in a transdisciplinary undergraduate course focused on real-world problem solving. *Int. J. Sustain. High. Educ.* **2013**, *14*, 404–433. [CrossRef]
18. Millis, B.J.; Cottell, P.G., Jr. *Cooperative Learning for Higher Education Faculty*; Series on Higher Education; Oryx Press: Phoenix, AZ, USA, 1997.
19. Bruffee, K.A. *Collaborative Learning: Higher Education, Interdependence, and the Authority of Knowledge*; Johns Hopkins University Press: Baltimore, MD, USA, 1993.
20. Rulifson, G.; Bielefeldt, A. Motivations to Leave Engineering: Through a Lens of Social Responsibility. *Eng. Stud.* **2017**, *9*, 222–248. [CrossRef]
21. Smith, J. Impacts of a humanitarian engineering education pathway on student learning and graduate outcomes. *Int. J. Serv. Learn. Eng. Humanit. Eng. Soc. Entrep.* **2019**, *14*, 1–20. [CrossRef]
22. Amadei, B.; Wallace, W.A. Engineering for humanitarian development. *IEEE Technol. Soc. Mag.* **2009**, *28*, 6–15. [CrossRef]
23. Wiek, A.; Withycombe, L.; Redman, C.L. Key competencies in sustainability: A reference framework for academic program development. *Sustain. Sci.* **2011**, *6*, 203–218. [CrossRef]
24. Cebrián, G.; Junyent, M. Competencies in education for sustainable development: Exploring the student teachers' views. *Sustainability* **2015**, *7*, 2768–2786. [CrossRef]
25. Sakhaee, E. The AAUL Framework of Leadership in Times of Crisis. In *Organization and Leadership in Disruptive Times*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 97–108.

26. Chaplin, G.; Wyton, P. Student engagement with sustainability: Understanding the value–action gap. *Int. J. Sustain. High. Educ.* **2014**, *15*, 404–417. [CrossRef]
27. Merritt, E.; Hale, A.; Archambault, L. Changes in Pre-Service Teachers' Values, Sense of Agency, Motivation and Consumption Practices: A Case Study of an Education for Sustainability Course. *Sustainability* **2018**, *11*, 155. [CrossRef]
28. Nousheen, A.; Zai, S.A.Y.; Waseem, M.; Khan, S.A. Education for sustainable development (ESD): Effects of sustainability education on pre-service teachers' attitude towards sustainable development (SD). *J. Clean. Prod.* **2020**, *250*, 119537. [CrossRef]
29. Yamanishi, K.; Li, H. Mining open answers in questionnaire data. *IEEE Intell. Syst.* **2002**, *17*, 58–63. [CrossRef]
30. Galhardi, L.B.; Brancher, J.D. Artificial Intelligence-IBERAMIA 2018. In Proceedings of the 16th Ibero-American Conference on AI, Trujillo, Peru, 13–16 November 2018; pp. 380–391. [CrossRef]
31. Mohler, M.; Bunescu, R.; Mihalcea, R. Learning to grade short answer questions using semantic similarity measures and dependency graph alignments. In Proceedings of the 49th Annual Meeting of the Association for Computational linguistics: Human Language Technologies, Portland, OR, USA, 19–24 June 2011; pp. 752–762.
32. Meurers, D.; Ziai, R.; Ott, N.; Kopp, J. Evaluating answers to reading comprehension questions in context: Results for German and the role of information structure. In Proceedings of the TextInfer 2011 Workshop on Textual Entailment, Edinburgh, UK, 30 July 2011; pp. 1–9.
33. Koufakou, A.; Gosselin, J.; Guo, D. Using Data Mining to Extract Knowledge from Student Evaluation Comments in Undergraduate Courses. In Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016; pp. 3138–3142. [CrossRef]
34. Abd-Elrahman, A.; Andreu, M.; Abbott, T. Using text data mining techniques for understanding free-style question answers in course evaluation forms. *Res. High. Educ. J.* **2010**, *9*, 1.
35. Bird, S.; Klein, E.; Loper, E. *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*; O'Reilly Media, Inc.: Newton, MA, USA, 2009.
36. Available online: https://spacy.io/models/en#en_core_web_sm (accessed on 31 July 2022).
37. Weischedel, R.; Palmer, M.; Marcus, M.; Hovy, E.; Pradhan, S.; Ramshaw, L.; Xue, N.; Taylor, A.; Kaufman, J.; Franchini, M.; et al. *Ontonotes Release 5.0 ldc2013t19*; Linguistic Data Consortium: Philadelphia, PA, USA, 2013; Volume 23.
38. Fellbaum, C. WordNet and wordnets. *Encyclopedia of Language and Linguistics*, 2nd ed.; Elsevier : Amsterdam, The Netherlands, 2005; pp. 665–670.
39. Peason, K. On lines and planes of closest fit to systems of point in space. *Philos. Mag.* **1901**, *2*, 559–572. [CrossRef]
40. Jones, K.S. A statistical interpretation of term specificity and its application in retrieval. *J. Doc.* **1972**, *28*, 11–21. [CrossRef]
41. Aizawa, A. An information-theoretic perspective of tf–idf measures. *Inf. Process. Manag.* **2003**, *39*, 45–65. [CrossRef]
42. Liu, J.; Chua, T.S. Building semantic perceptron net for topic spotting. In Proceedings of the 39th Annual Meeting on Association for Computational Linguistics-ACL '01, Toulouse, France, 6–11 July 2001; pp. 378–385. [CrossRef]
43. Dillon, J.; Bosch, N.; Chetlur, M.; Wanigasekara, N.; Ambrose, G.A.; Sengupta, B.; D'Mello, S.K. Student Emotion, Co-Occurrence, and Dropout in a MOOC Context. *Int. Educ. Data Min. Soc.* **2016** .
44. White, A.E.; Moeller, J.; Ivcevic, Z.; Brackett, M.A. Gender identity and sexual identity labels used by US high school students: A co-occurrence network analysis. *Psychol. Sex. Orientat. Gend. Divers.* **2018**, *5*, 243. [CrossRef]
45. Leydesdorff, L.; Nerghes, A. Co-word maps and topic modeling: A comparison using small and medium-sized corpora (N < 1,000). *J. Assoc. Inf. Sci. Technol.* **2017**, *68*, 1024–1035. [CrossRef]
46. Rude, S.; Gortner, E.M.; Pennebaker, J. Language use of depressed and depression-vulnerable college students. *Cogn. Emot.* **2004**, *18*, 1121–1133. [CrossRef]
47. Robinson, R.L.; Navea, R.; Ickes, W. Predicting Final Course Performance From Students' Written Self-Introductions. *J. Lang. Soc. Psychol.* **2013**, *32*, 469–479. [CrossRef]
48. Syah, T.; Nurhayaty, A.; Apriyanto, S. Computerized Text Analysis on Self-Description Text to Get Student's Prevailing, Confidence, and Drives. *J. Phys. Conf. Ser.* **2021**, *1764*, 012056. [CrossRef]
49. Pennebaker, J.W. Putting stress into words: Health, linguistic, and therapeutic implications. *Behav. Res. Ther.* **1993**, *31*, 539–548. [CrossRef]
50. Pennebaker, J.W.; Francis, M.E.; Booth, R.J. Linguistic inquiry and word count: LIWC 2001. *Mahway Lawrence Erlbaum Assoc.* **2001**, *71*, 2001.
51. Pennebaker, J.W.; Booth, R.J.; Francis, M.E. LIWC2007: Linguistic inquiry and word count. *Austin Tex. Liwc. Net* **2007**.
52. Pennebaker, J.W.; Booth, R.J.; Boyd, R.L.; Francis, M.E. *Linguistic Inquiry and Word Count: LIWC2015*; Pennebaker Conglomerates: Austin, TX, USA, 2015. Available online: www.LIWC.net (accessed on 31 July 2022).
53. Watson, M.K.; Pelkey, J.; Noyes, C.; Rodgers, M.O. Using Kolb's learning cycle to improve student sustainability knowledge. *Sustainability* **2019**, *11*, 4602. [CrossRef]
54. Mehl, M.R.; Pennebaker, J.W. The Sounds of Social Life: A Psychometric Analysis of Students' Daily Social Environments and Natural Conversations. *J. Personal. Soc. Psychol.* **2003**, *84*, 857–870. [CrossRef]
55. Schnurr, P.P.; Rosenberg, S.D.; Oxman, T.E.; Tucker, G.J. A Methodological Note on Content Analysis: Estimates of Reliability. *J. Personal. Assess.* **1986**, *50*, 601–609. [CrossRef]

sustainability

MDPI

Article

Bibliometric Literature Review of Adaptive Learning Systems

Dionisios Koutsantonis [1], Konstantinos Koutsantonis [1], Nikolaos P. Bakas [1,*], Vagelis Plevris [2], Andreas Langousis [3] and Savvas A. Chatzichristofis [4,*]

[1] Research and Development Department, RDC Informatics, 10442 Athens, Greece
[2] Department of Civil and Architectural Engineering, Qatar University, Doha P.O. Box 2713, Qatar
[3] Department of Civil Engineering, University of Patras, 26504 Patras, Greece
[4] Intelligent Systems Laboratory, Department of Computer Science, Neapolis University Pafos, Pafos 8042, Cyprus
* Correspondence: n.bakas@rdc.gr (N.P.B.); s.chatzichristofis@nup.ac.cy (S.A.C.)

Abstract: In this review paper, we computationally analyze a vast volume of published articles in the field of Adaptive Learning, as obtained by the Scopus Database. Particularly, we use a query with search terms targeting the area of Adaptive Learning Systems by utilizing a combination of specific keywords. Accordingly, we apply a multidimensional scaling algorithm to construct bibliometric maps for keywords, authors, and references. Subsequently, we present the computational results for the studied dataset, reveal significant patterns appearing in the field of adaptive learning and the inter-item associations, and interpret the findings based on the current state-of-the-art literature in the area. Furthermore, we demonstrate the time-series of the evolution of the research terms, their trends over time, as well as their prevalent statistical associations.

Keywords: adaptive learning; intelligent tutoring systems; personalized learning; machine learning; bibliometrics

Citation: Koutsantonis, D.; Koutsantonis, K.; Bakas, N.P.; Plevris, V.; Langousis, A.; Chatzichristofis, S.A. Bibliometric Literature Review of Adaptive Learning Systems. *Sustainability* **2022**, *14*, 12684. https://doi.org/10.3390/su141912684

Academic Editor: Jian-Hong Ye

Received: 6 August 2022
Accepted: 28 September 2022
Published: 5 October 2022

1. Introduction

Scientific research in the field of learning systems has progressed through the inevitable utilization of technology advancements [1]. The aim is to describe and design learning processes that are able to dynamically adjust to the level or type of course content based on an individual student's tacit knowledge or skills. Accordingly, systems that accelerate a learner's performance have emerged, often characterized as intelligent tutoring systems and personalized training [2]. Among others, these terms have been referenced in numerous scientific papers over the last 20 years, to describe the notion and aspects of adaptive learning systems. However, due to the vast scientific output produced in any discipline, it is a modern approach to use quantitative tools to analyze the current structure of a thematic area [3], usually followed by a more specific analysis of the most influential papers [4].

The purpose of this paper is to systematically analyze the evolution, status, and trends of scientific research in the field of adaptive learning systems, by utilizing a generic framework for Bibliometric Analysis [5–8]. This analysis is performed by assorting specific attributes of the papers, such as the comprised keywords, authors, references, as well as other types of bibliographical information, e.g., journal's metrics and papers' citations, herein called Bibliometric Objects (BO), and using a machine learning algorithm, which is described in Section 2. In Section 3, the query for the extraction of the studied datasets is described, and in Section 4 we analyze the numerical results. In Section 5 we associate the results with the top-cited papers in the field and interpret the findings. In Section 7, the conclusions of the study are presented.

2. Multidimensional Scaling for Bibliometric Mapping

The rendering of a bibliometric map requires the identification of the location of each BO on a two-dimensional space, such that the distances among the objects represent the inter-item dissimilarities. Particularly, a bibliometric map has the following characteristic attributes:

- Each object (e.g., Keyword, Author, or Reference) is represented as a point on the 2D map, with its coordinates on the Cartesian plane;
- The objects with co-occurrences are connected with a line;
- The thickness of the line represents the link strength, which is proportional to the similarity (or co-occurrence) between the objects;
- The distances between the objects are indicators of their dissimilarity.

2.1. Baseline Formulation

We define as $\mathbf{c} = c_{ij}$ the elements of the contingency matrix $\mathbf{c}$, with

$$[N] = \{1, 2, \ldots, N\},$$

the iterator for the number of Bibliometric Objects N, and $i, j \in [N]$. The contingency table c_{ij} corresponds to the counting of the co-occurrence of objects. Henceforth, it is rational to define the similarity matrix $\mathbf{s}$, with elements

$$s_{ij} := \frac{c_{ij}}{\max \mathbf{c}}, \tag{1}$$

and the corresponding dissimilarity

$$ds_{ij} := 1 - s_{ij}, \tag{2}$$

both comprising values in $[0, 1]$. Accordingly, we define the pairwise distances of two elements on the map

$$d_{ij} := \|\mathbf{x}_i - \mathbf{x}_j\|_2^2, \tag{3}$$

where $\mathbf{x}_i$ stands for the t-dimensional vector, defining the position of the *ith* point on the map. For the case of two-dimensional maps, we use $t = 2$; however, the procedure can be generalized for $t > 2$ as well for 3D maps or dimensionality reduction procedures.

Ideally, we should have a map such that

$$ds_{ij} \equiv d_{ij} \quad \forall i, j \in [N], \tag{4}$$

or alternatively,

$$ds_{ij} \propto d_{ij} \quad \forall i, j \in [N]. \tag{5}$$

However, this cannot be attained exactly for any $N > 3$. Henceforth, the problem of finding the bibliometric objects' location on the map should be formulated as an optimization problem, that is, to find the best possible Furthermore, the formulation of a representative objective function for such a purpose is not unique and vastly affects the obtained shape of the map (see next Section 2.2). Moreover, the optimization algorithm used to solve this problem is still an open issue (see below Section 2.3).

2.2. Objective Functions

By defining f as the objective function, the problem of finding the best possible topology of the objects is formulated as finding the optimal $\mathbf{x}$, that is to say

$$\begin{aligned} \mathbf{x}_{opti} = \arg\min f(\mathbf{x}) := \{ & \mathbf{x} \in A \subseteq \mathbb{R}^t \\ & \mid \forall \mathbf{y} \in A : f(\mathbf{y}) \geq f(\mathbf{x}) \}, \end{aligned} \tag{6}$$

for a t-dimensional map, and a given domain $A \subseteq \mathbb{R}^t$. A number of objective functions were investigated in order to construct a map, such that each d_{ij} corresponds to the reciprocal ds_{ij}. One approach to this multidimensional scaling problem [9], can be formulated as finding a t-dimensional map, where the order of distances of the N objects, when ranked monotonically at a descending order, corresponds to the monotonically descending order of their dissimilarities [10,11]. That is to say, higher distances resemble lower similarities. This is achieved through the minimization of the corresponding errors, called stresses. However, this approach does not always offer a consistent representation of the dissimilarities, as it does not utilize a mathematical association for the mapping of the dissimilarities to the distances further than the monotonicity. Thus, two pairs of points on the map with comparable distances may correspond to dissimilarities with a high difference.

Another approach often used in bibliometric mapping [12], is to minimize the objective function

$$f(\mathbf{x}) = \sum_{i,j \in [N]} s_{ij} d_{ij}, \tag{7}$$

or, equivalently,

$$f(\mathbf{x}) = \sum_{i,j \in [N]} \frac{d_{ij}}{ds_{ij}}. \tag{8}$$

However, this approach does not exploit the significant information where the similarities are equal to zero because they will be omitted in the summation. Additionally, it does not offer a measure of the efficiency of the algorithm, as the summation is not comprehensible.

Additionally, we investigated as an objective function a combined metric obtained by fitting a regression curve between the dissimilarities and the distances. The regression was nonlinear, considering the dissimilarities to an *rth* power as the independent variable and the corresponding distances as the dependent ones. The regression equation is written as

$$d_{ij} = \alpha ds_{ij}^{r} + \beta + \varepsilon_{ij},$$

where α is the regression coefficient, β is the constant term, and ε is the regression residuals. Hence, by varying the positions on map $\mathbf{x}_i$, the regression coefficients a, b, as well as the coefficient of determination R^2, also varied. Accordingly, the objective function to be minimized can be written as

$$f(\mathbf{x}) = |\beta(\mathbf{x})| + \left|1 - R(\mathbf{x})^2\right| + \frac{|\alpha(\mathbf{x}) - 1|}{3},$$

using the $\frac{1}{3}$ factor for α, for normalization purposes. This way, the optimization algorithm is utilized to find a solution satisfying the fundamental regression requirements, for low or zero constant term, high R^2 and close to a unit slope of the regression line. The total number of design variables was $2N + 1$; i.e., $2N$ for the points and 1 for r, as we varied r during the optimization process. The regression coefficient α that was divided by three relaxed the variation of the slope, as it is important to avoid extreme values, for example, slopes close to zero or tending to infinity. However, this formulation is in fact a double optimization problem, one for the map and one for the regression, which is nested into the former. Hence, the execution time was very slow. Additionally, the results were not satisfactory, as the errors, although small, disorientated the mapping procedure.

Accordingly, another idea was to minimize directly, the absolute value of the difference between the distances, and a nonlinear function of the dissimilarities

$$f(\mathbf{x}, \alpha, \beta, r) = \left|\alpha + \beta ds_{ij}{}^{r} - d(\mathbf{x})_{ij}\right|.$$

Therefore, by including α, β in the optimization problem, the number of variables becomes $2N + 3$. However, this approach gave us good results only for extreme (high or small) values of α, β; thus, the corresponding maps did not supply any information regarding the efficiency of the algorithm.

Additional approaches were investigated. In particular, instead of using the co-occurrence map, the matrix of the chi-squared statistic, as well as the corresponding p-values, were investigated as candidates for similarity matrices. However, the p-values exhibit vastly low values (e.g., 10^{-100} or less), and thus the corresponding distances could not be coherent. Similarly, for the χ^2, the values were either equal to zero or close to 1, not supplying the necessary information to construct the map. Furthermore, instead of minimizing the summation of the errors between the distances and the dissimilarities, the maximum difference, as well as specific quantiles of the distribution of the differences, were also investigated. However, they did not perform well enough, as the minimization problem does not have a solution equal to zero because of the error always existing in dimensionality reduction problems. Thus, optimizing the maximum or the 75% percentile does not necessarily lead to the minimization of the other distances, as well.

Another idea was also investigated, inspired by the dropout technique [13], which is common in the training of artificial neural networks, targeting the avoidance of overfitting. In particular, the technique was used each time the optimizer did not optimize the a% percentile (i.e., 75%) but the $a + \varepsilon$, where ε stands for a randomly generated error number. However, this technique did not give satisfactory results for the reasons stated above. Other specific techniques, such as linkage or hierarchical clustering, exist for the representation of multiple correlations between N objects; however, they do not usually offer a coherent, easy, and quick interpretable way of representing the associations between N objects simultaneously, as that multidimensional scaling offers.

Finally, after many trials, we decided to utilize the formula

$$f(\mathbf{x}) = \frac{1}{N} \sum_{i,j \in [N]} |ds_{ij} - d_{ij}|. \tag{9}$$

We may also add the constraint

$$\min d_{ij} \geq l, \tag{10}$$

where l is a specified threshold to avoid the coincidence of pairs of objects in the final map.

Furthermore, we may easily weigh the elements of the objective function by multiplying them by the summation of the occurrences for each pair i, j. This way, we can control the map creation by giving more attention to the most important elements.

2.3. Optimization Algorithms

A number of optimization algorithms were investigated for the minimization of the objective function of the optimization problem, including genetic algorithms [14], pattern search [15], particle swarm optimization [16–18], trust region [19], global search [20], and the Levenberg–Marquardt Algorithm [10,21,22], among others. For the trust region and global search methods, the supply of the gradient was investigated, as well. Finally, we utilized the Pure Random Orthogonal Search (PROS) optimization Algorithm [23,24], modified to fit the Bibliometric problem, as presented in Algorithm 1.

Algorithm 1: Bibliometric map generation

Data: Vector of Strings of the Bibliometric Objects
Result: optimal positions $\mathbf{x}_{opti} = \mathbf{x}_{i,opti} \forall i \in [N]$ on the Bibliometric Map
Compute co-occurrences c_{ij} of the studied BO and maximum iterations f_e;
Compute s, ds from c_{ij} (Equations (1) and (2));
Initialize $\mathbf{x} := \mathbf{x}_{i \in [N]} \in A$ randomly;
Assign $\mathbf{x}_{opti} \leftarrow \mathbf{x}$;
Compute $f_{opti} = f(\mathbf{x}_{opti})$;
for $i \in f_e$ **do**
  $i \leftarrow \mathcal{U}(1, N)$;
  $j \leftarrow \mathcal{U}(1, 2)$;
  $ra \leftarrow \mathcal{U}(0, 1)$;
  $x_{ij} = lb_j + ra \times (ub_j - lb_j)$;
  Compute $d_{ij} \quad \forall \quad i, j \in [N]$;
  Compute $f_{run} = f(\mathbf{x})$ (Equation 9);
  if $f_{run} \leq f_{opti}$ **then**
    $f_{opti} \leftarrow f_{run}$;
    $\mathbf{x}_{opti} \leftarrow \mathbf{x}$;
  else
    $\mathbf{x} \leftarrow \mathbf{x}_{opti}$;
Plot Bibliometric Map;
return $\mathbf{x}_{opti}$ positions

The procedure is generic for all BO, denoting either keywords, authors, references, or any other object. All the matrices $\mathbf{s}, \mathbf{c}, \mathbf{ds}, \mathbf{d}$ are symmetric. The vector $\mathbf{x}$ comprises the design variables of the optimization problem. For a 2D map, $\mathbf{x}$ denotes a set of points in the $\mathbf{R}^2$ space, that is

$$\mathbf{x}_{i \in [N]} = (x_{i1}, x_{i2}). \tag{11}$$

Ideally and only theoretically, all the errors $e_{ij} = ds_{ij} - d_{ij}$ should become equal to zero, but this cannot be attained in practice because the map represents multidimensional relationships in the 2D space, leading in dimensionality reduction. Hence, the aim is to find all the $\mathbf{x}_i$, corresponding to the minimum errors e_{ij}.

The algorithm starts with the initial calculation of the co-occurrence table of the studied objects (keywords, authors, or references). Accordingly, the similarity and dissimilarity matrices are obtained, as per Section 2.1. The optimization algorithm randomly initializes the positions of each object and computes the distances between the objects on the bibliometric map. In the end, the optimal values of the positions are utilized to visualize the results on the bibliometric map.

Furthermore, we may specify a priori the locations of the top frequent keywords $[t] = \{1, 2, \ldots, t\}, \mathbf{x}_{[t]}$. For this work, we positioned the most frequent keyword at the center of the map (0, 0) and the next four most frequent on a unit rectangle around zero. The optimization algorithm runs after this step accordingly.

3. The Studied Databases of Papers

A bibliometric study depends on the papers' database; subsequently, the initial effort is given to assembling a thorough query. Some iterations might be necessary until the query represents the actual topic studied. It is essential to keep track of the searching date, as well as the exact query, in order to obtain reproducible results. A good practice is to start with a generic search and, step-by-step, narrow it down to identify the thematic area in the database in order to obtain a final search query that will focus on a particular field without losing its significant thematic areas. There are several databases that can be used such as Scopus (https://www.scopus.com, (accessed on 17 March 2022)), Web of

Science (https://www.webofscience.com, (accessed on 17 March 2022)), Google Scholar (https://scholar.google.com/, (accessed on 17 March 2022)), etc. In this work we use the Scopus database, comprising a rich collection of journals [25].

3.1. The Scopus Query

We searched in the Scopus [26] database by using a combination of keywords derived from an initial screening of the papers in the literature regarding Adaptive Learning (Table 1). Hence we get the following query of the form "**TITLE-ABS-KEY ("personalized learning") OR TITLE-ABS-KEY ("intelligent tutoring systems")** . . . ", etc. The complete structure of the query is presented in Table 1. We selected these particular keywords, after an overview of the literature regarding adaptive learning systems. Accordingly, after a manual search of the existing papers, we selected the ones that occurred with high frequency, still relevant to the studied topic.

Table 1. Search query in Scopus.

Adaptive Educational Hypermedia System	**OR**	Adaptive Educational System	**OR**
adaptive learning	**OR**	advanced learning technologies	**OR**
intelligent learning platforms	**OR**	intelligent tutoring systems	**OR**
AI-based learning systems	**OR**	personal learning environments	**OR**
personalized learning	**OR**	smart learning environments	**OR**
tutor-based expert systems	**OR**	web-based adaptive educational applications	
	AND		
education OR lesson	in	Title OR Keywords	

We assembled the query on 17 March 2022. We limit the results to the last 30 years, i.e., within the period 1992–2021, as before 1992 we had only a few papers (less than 20 per year), and 2022 is still a year in progress. Hence, we obtained a total of 5564 papers in the following categories: Conference Papers (N = 3756), Articles (N = 1489), Book Chapters (N = 129), Conference Reviews (N = 75), Reviews (N = 67), Books (N = 19), Notes (N = 11), Retracted Articles (N = 6), Editorials (N = 6), Letters (N = 3), Errata (N = 2), Short Surveys (N = 1).

3.2. Keywords' Grouping

Keywords' grouping is a significant part of the analysis, as some keywords appear with multiple records. For example, the keyword "Artificial Neural Network", will count as different than "Artificial Neural Networks", and "ANNs". The grouping of keywords after a first analysis will contribute to obtaining aggregated and meaningful results. In Appendix A, we provide the grouping of the utilized keywords . All keywords (as well as authors, references, etc.) are converted to lowercase so as to avoid duplicate BO.

4. Results

4.1. Bibliometric Map of Keywords

In Figure 1, the bibliometric map of the top 28 keywords after grouping is presented. As stated in Section 2, we put in the center the keyword "**intelligent tutoring system**", and around it the next top four keywords, which are:

- Adaptive learning;
- Personalized learning;
- Artificial intelligence;
- Higher education.

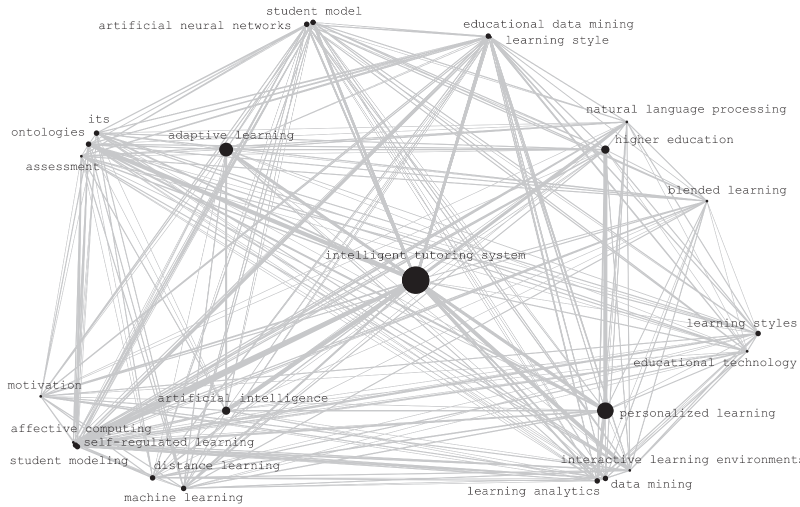


Figure 1. Association map of top 28 keywords, for adaptive learning database, after their grouping.

Despite the complex nature of the associations presented by the lines connecting the keywords, we may distinguish some clusters of keywords as automatically detected by the algorithm. Particularly, close to keyword "**artificial intelligence**", we observe the following keywords:

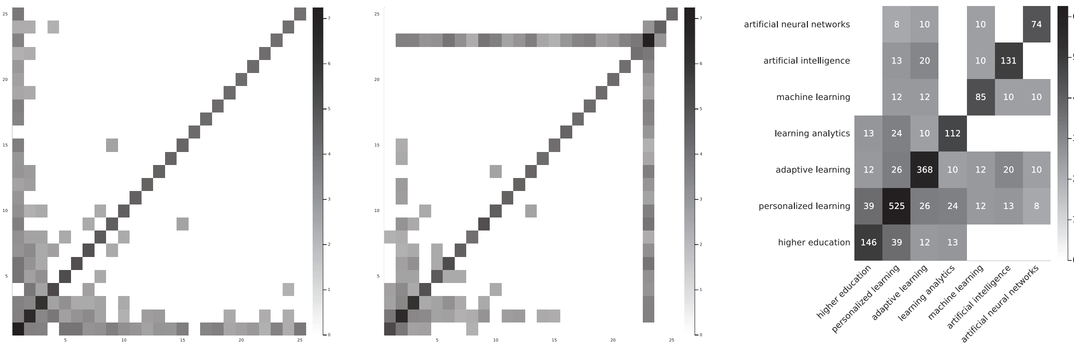
- Self-regulated learning;
- Affective computing;
- Machine learning;
- Distance learning;
- Student modeling,
- Motivation.

Furthermore, close to "**adaptive learning**", we see "ontologies", and "assessment". Additionally, close to "**personalized learning**", we observe the keywords "interactive learning environments", "data mining", "learning analytics" at a micro cluster, along with keywords "educational technology", and "learning styles" at another cluster near the former one. The keyword "**higher education**", was found close to "natural language processing", and "blended learning". Finally, at the top edge of the map, we may see the keyword "student model", very close to "artificial neural networks", and "learning style", close to "educational data mining". Accordingly, we present in Table 2 the automatically extracted sub-categories of the field.

Table 2. Sub-categories of adaptive learning systems research.

Artificial Intelligence	Adaptive Learning	Personalized Learning	Higher Education
self-regulated learning	ontologies	interactive learning environments	natural language processing
affective computing	assessment	data mining	blended learning
machine learning		learning analytics	
distance learning		educational technology	
student modeling		learning styles	
motivation			

Similar findings are rendered if we re-arrange the co-occurrence matrix for keywords, using the Cuthill McKee algorithm [27,28] as per Figure 2. Particularly, by re-arranging the rows and columns of the matrix to obtain the majority of the elements near the main diagonal, we see that the keywords "higher education", "personalized learning", "adaptive learning", "learning analytics", "machine learning", "artificial intelligence", and "artificial neural networks" are highly related pair-wise.

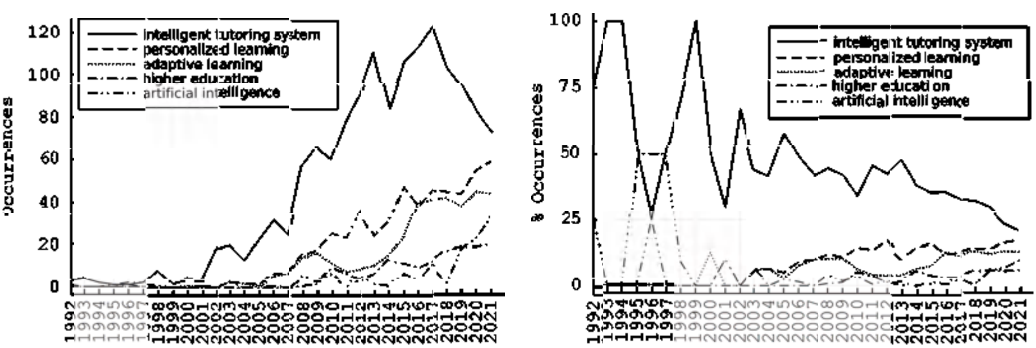


(**a**) Initial Contingency (**b**) CuthillMcKee Contingency (**c**) CuthillMcKee Matrix—Center

Figure 2. Rearrangement of the contingency matrix using the CuthillMcKee method.

4.2. Current State and Emerging Insights

It is of great importance to highlight not only the structure of the studied field and its sub-categories, but also the evolution of the research, along with future trends. In order to do this in a generic, quantitative manner, we perform time-series analysis. Particularly, in Figure 3a, we demonstrate the evolution of the number of the top five keywords' occurrences over time. We may observe a clear increasing pattern for all keywords, apart from "intelligent tutoring system", which presents a decreasing pattern after 2017. This is further confirmed if we plot the normalized time-series, which is the initial counting, divided by the total number of papers published in the current year for the specific data set. This is depicted in Figure 3b, where "intelligent tutoring system", decreases constantly, contradictory with the other top four keywords.



(**a**) Occurrences (**b**) Normalized Occurrences

Figure 3. Time series of the evolution of the top keywords.

Furthermore, by identifying all the pairwise Pearson correlation coefficients of the time-series, we obtain Figure 4, depicting "learning styles" to be highly correlated over time with "assessment", with a peak in 2001, a Pearson correlation of 92.7%, and, similarly "learning styles" with "artificial neural networks", having a Pearson correlation of 79.3%.

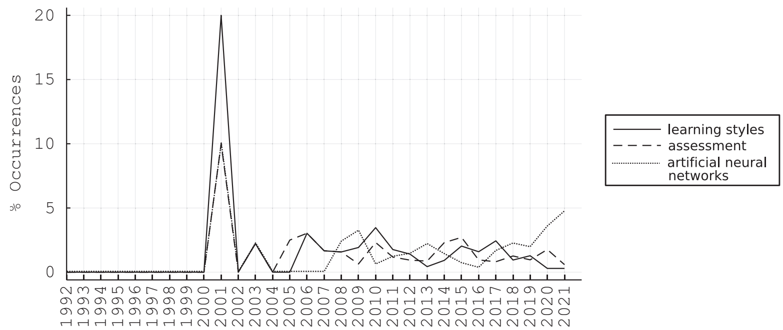


Figure 4. Computational identification of correlation among "learning styles", "assessment", and "artificial neural networks" timeseries.

An interesting finding is obtained from Figure 5, where we may see that "educational data mining" and "natural language processing" time-series exhibit close evolution over time, with a Pearson correlation of 85.1%. This indicates a significant part of the modern adaptive learning system, where text analysis is prevalent.

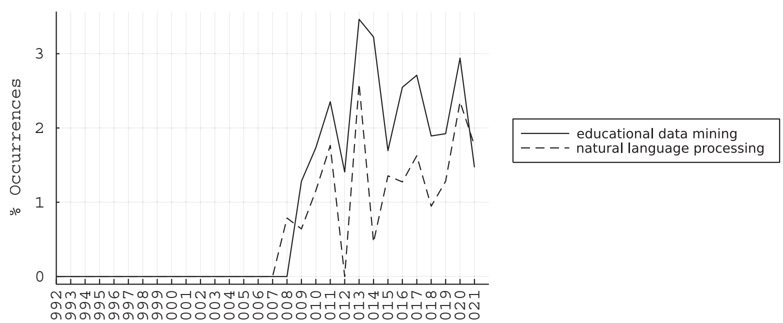


Figure 5. Similar evolution of "educational data mining" and "natural language processing" time series.

Finally, in Figure 6, we automatically detect the five top nonlinear trends by fitting a second order polynomial in the time-series, and selecting the ones with higher coefficients.

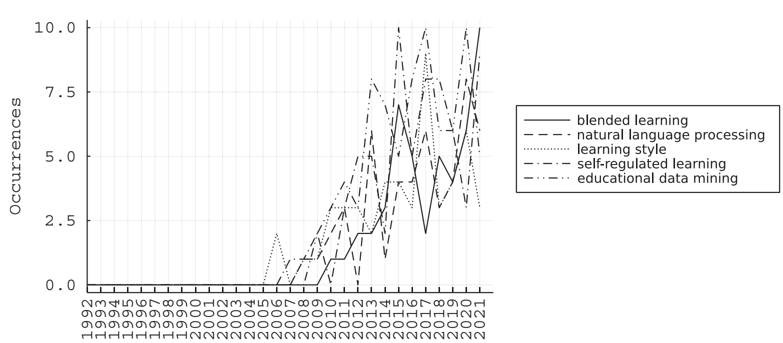


Figure 6. Time series of maximum nonlinear trends.

Hence, we identify as emerging trends the following keywords: "blended learning", "natural language processing", "learning style", "self-regulated learning", and "educational data mining".

4.3. Map of References

In total, we identified 131,898 references in the studied papers. The top 25 papers cited within the studied database were extracted to create their bibliometric map. Four of them did not appear simultaneously as references; hence 21 were finally kept [29–42] (Figure 7). Accordingly, we observe four distinct clusters of references.

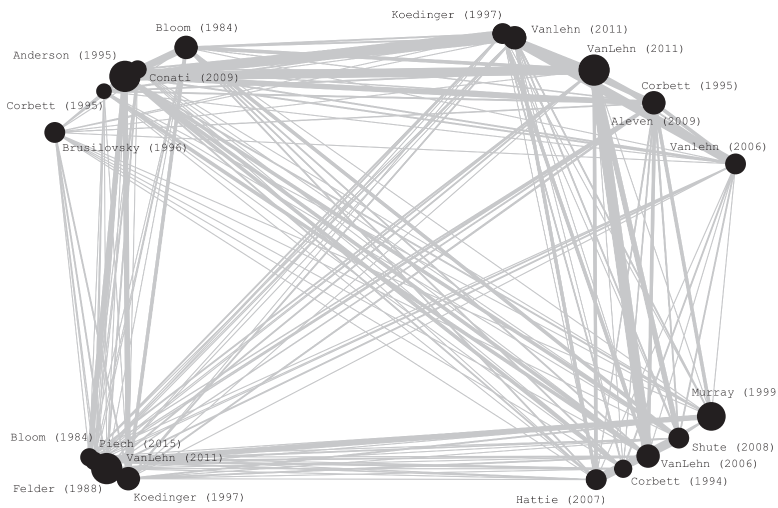


Figure 7. Association map of references [29–42].

5. Analysis of Top Cited Papers

5.1. Highly Cited Original Research Articles

This section revisits and briefly presents the contribution of the 10 most cited original research articles retrieved using the query described in Section 3. For each manuscript, we attempt to identify how its content and scope relate to the terms listed on the top-keywords map.

The authors in Ref. [43] performed a comparative study between an inverted classroom and the traditional classroom for an introductory statistics course at the same university. The findings of this study demonstrated that students participating in the inverted classroom were less satisfied with the teaching format than students in the traditional classroom were. Yet, they became more open to cooperative learning and innovative teaching methods. Although the inverted classroom is not explicitly referred to in the map, it is however directly related to the term "blended learning" as the inverted classroom is a specific type of blended learning.

Baker et al. [44] examined student nonverbal behaviors such as boredom, frustration, confusion, and engagement during interactions with three computer-based learning environments, including a dialogue tutor and a problem-solving-based intelligent tutoring system. The results showed that boredom in computer learning environments is associated with poorer learning and problem behavior, while frustration was less associated with more inadequate understanding. In this case, we can easily correlate the term intelligent tutoring system with the keyword map as it appears to be the map's core term.

Another study that refers to the term "intelligent tutoring systems" of the map is the work of VanLehn et al. [45]. Their research reported results from 5 years of using the Andes [45,46] tutoring system in introductory physics courses taught at the U.S. Naval Academy. The authors noted that students who used the Andes for homework got significantly higher exam scores than students in control groups who did paper-and-pencil homework.

In a similar study [47], the researchers used a probabilistic student model with Bayesian networks in the Andes tutoring system to manage uncertainty. This intelligent tutoring system hints at "rules" when it infers that a student has not yet mastered the lesson. Regarding the keyword map, besides the "intelligent tutoring system" term, the term "machine learning" also appears as the authors performed both a machine learning style evaluation and an evaluation with real students of the Andes student model for problem-solving.

In addition, Chen et al. [48] proposed a personalized e-learning system based on IRT (PEL-IRT), which considers both the course material's difficulty and the learner's ability to provide individual learning paths for learners. Experimental results show that the proposed system can provide personalized learning and help learners learn more effectively and efficiently. The term "personalized learning", which is the researchers' primary goal, can be found among the keywords of the map.

In Ref. [49], the authors present a web-based Adaptive Educational Hypermedia system, called INSPIRE, designed to support web-based personalized instruction and traditional classroom-based teaching as a supplementary resource. The proposed method includes adaptive and adaptable behavior guided by the learner model and characteristics such as knowledge level and learning style. Regarding the keyword map, we can easily recognize the "adaptive learning", "personalized learning", "learning styles", and "student model" terms.

ELM-ART [50] is another adaptive remote tutor system implemented on WWW that combines the features of an electronic textbook, a learning environment, and an intelligent tutoring system, helping students learn to program in Lisp. The system supports the student by applying adaptive hypermedia techniques—adaptive annotation and adaptive link sorting. As before, "intelligent tutoring system", and "adaptive learning" are among the terms in the keyword map.

In another study, Chen and Chun [51] presented an analysis of a personalized mobile English vocabulary learning system in mobile devices providing appropriate English vocabulary for learning. Results showed that the vocabulary learning system promoted the students' performance and interest because of the practical and flexible vocabulary learning process. According to most of the learners who participated in the experiments, the proposed system was very helpful in English vocabulary learning. It promoted their learning interests and English vocabulary abilities without time or constraints according to their vocabulary ability and memory cycle. Furthermore, by observing the terms of the keyword map, we can indirectly associate the terms "distance learning" and "learning styles" terms with the paper.

Vrugt et al. [52] introduced a self-adaptive multimethod evolutionary search concept and proposed the multi-algorithm genetically adaptive method (AMALGAM). The AMALGAM-SO merges different optimization methods such as covariance matrix adaptation evolutionary strategy, parental-centric recombination operator, PSO, GA, and DE for population evolution [17]. The optimization methods with the highly productive search were adaptively selected to generate more offspring during the investigation. Experimental results conducted using a set of standard benchmark functions showed that the concept of self-adaptive multimethod optimization improved the robustness and efficiency of the evolutionary search. For this paper, we can see that the term "genetic algorithm" is not explicitly referred to in the map as a keyword but is directly related to the terms "machine learning" and "artificial intelligence".

Finally, in Ref. [53], researchers proposed a three-interval type-2 fuzzy neural network (IT2FNN) architecture, with hybrid learning algorithm techniques (gradient descent backpropagation and gradient descent with adaptive learning rate backpropagation), for solving the uncertainty of real-life problems. Through the two conducted experiments (i.e., a nonlinear identification problem for control systems simulation and a nonlinear Mackey–Glass chaotic time series prediction problem with uncertainty sources), the efficiency of the IT2FNN algorithm was validated. "Artificial neural networks", "artificial intelligence", and "machine learning" terms on the map are directly or indirectly related to Ref. [53].

5.2. Highly Cited Original Research Articles Published in the Last 5 Years

This section continues by presenting the most highly cited articles published in the last five years in order to discuss recent significant advancements.

The authors in Ref. [54] analyzed the implementation of adaptive learning systems in the learning process of higher education according to the scope, type of adaptive learning, functional purpose, integration with existing learning management systems, and application of modern technologies. Researchers support that utilizing the blended learning concept makes it easier to structure an e-learning course and define how it can be adapted. By observing the terms of the keyword map, we can easily recognize the terms "blended learning", "adaptive learning", and "higher education".

Ullah et al. [55] proposed an expanded Technology Acceptance Model for implementing blockchain that could assist decision-makers in building an intelligent learning environment for the educational institutes of the emerging economies. Although no specific terms can be directly identified in the keyword map, the term "adaptive learning" can indirectly be related to the paper.

In addition, in their study, Molenaar et al. [56] investigated the student's self-regulated learning (i.e., how learners regulate their effort, accuracy, and education) with adaptive learning technology using moment-by-moment learning curves. Their result confirmed that adaptive technology could improve self-regulated learning and increase learning efficiency. In this case, "self-regulated learning", "adaptive learning", and "learning analytics" terms can be directly and indirectly associated with the keyword map.

Han et al. [57] used learning analytics to develop two dashboards for students and instructors. This enabled students to monitor their face-to-face collaborative argumentation (FCA) process through adaptive feedback and helped the instructor provide adaptive support at the right time. The paper easily relates to the keyword map's terms "learning analytics" and "adaptive learning".

Furthermore, in Ref. [58], researchers conducted an integrative literature review in higher education to identify the most applicable and effective technology models that support personalized learning within blended learning environments. The study highlighted three models:

- Digital badges;
- A learner dashboard as the main feature and adaptive learning technology;
- Competency-based technology that adopts algorithm-based tutoring systems.

The terms "blended learning", "personalized learning", and "higher education", which are the researchers' focus, also appear in the keyword map.

Shiff [59], in his recent work, assessed the status of artificial intelligence in education (AIEd's), with a focus on intelligent tutoring systems and anthropomorphized artificial teaching agents. According to the author, the most prominent AIEd characteristics are personalized real-time feedback and adaptive learning. Correlating this paper with the keyword map, one can easily recognize the terms "artificial intelligence", "adaptive learning", and "personalized learning".

In Ref. [60], the prototype for a fully automated gaze-based attention-aware adaptive intelligent tutoring system for classroom use is presented. To do so, Hutt et al. [60] leverage eye trackers and pre-trained machine learning models to detect mind-wandering, using data from high-school students. At the same time, they interacted with educational technology in classrooms. The researchers used this approach to trigger dynamic interventions to reengage attention and improve learning. The paper is, directly and indirectly, related to the terms "educational technology", "machine learning", "adaptive learning", and "artificial intelligence" of the keyword map.

Tabuenca et al., in their work [61], conducted a systematic literature review to investigate an intelligent learning environment (SLE) according to three parameters:

1. What affordances make a learning environment smart?
2. Which technologies are used in SLEs?
3. In what pedagogical contexts are SLEs used?

According to the authors, smart learning can be analyzed as learning "in interactive, intelligent, and tailored environments, supported by advanced digital technologies and services". Based on this definition, we can easily correlate the paper with the terms "adaptive learning", "personalized learning", "interactive learning environments", and "educational technology".

Finally, Christopoulos and Sprangers [62] studied the factors affecting the perspective of teachers and students on ICT integration for mathematics teaching during the COVID-19 pandemic. Among the factors examined in their paper are:

- Different contexts (school vs. home) and circumstances (in-person vs. remote learning);
- Demographics;
- Confidence in using technology;
- Perspective on technical usefulness;
- Platform evaluation.

The correlation of Ref. [62] with the terms in the keyword map is achieved through the words "educational technology", and "distance learning".

6. Discussion

The analysis of the articles reinforces the conclusions drawn from the study of the co-occurrence of keywords, as discussed in Section 4. Both older and modern literature focuses on adaptive learning and personalized learning methodologies with an emphasis on using artificial intelligence. In recent years, however, particular emphasis has been placed on developing blended and distanced learning systems. This is to be expected, given that the COVID-19 pandemic has forced the education systems to adapt to the new reality globally.

We now find the terms "educational data mining" and "natural language processing" co-occurring in modern literature. The educational data mining process focuses on creating techniques for examining the distinctive, massive amounts of data generated in educational settings and using them to better understand students and the environments in which they learn. Today, the literature offers a large amount of data on student behavior, allowing researchers to conduct in-depth analyses and draw valuable conclusions. In addition, the significant development of artificial intelligence techniques provides researchers with tools for accurate and efficient automated data analysis without requiring special technical knowledge.

It is also important to note that the current literature focuses on self-regulated learning techniques. Self-regulated learning is described as the ability of a learner to understand and control his/her learning environment. The students are asked to define their goals and search for self-learning and self-reinforcement methods. Intelligent tutoring systems are called upon to support the student's effort, monitor his/her progress, and suggest methods to improve the process.

Observing the technologies used in adaptive learning systems, it is clear that the recent literature focuses on adopting artificial neural networks, leaving behind other artificial intelligence approaches. Thus, the term "intelligent tutoring systems" is being less utilized, and the current research focuses on "personalized" and "adaptive" learning, with the usage of machine learning algorithms. This is, of course, to be expected, given that this technology seems to dominate the field of artificial intelligence. The fact that artificial neural networks are nowadays easy to use plays an important role in shaping effective and efficient adaptive learning systems. Today, many open-source solutions allow their adoption by researchers with minimal technological background. As a result, educational researchers have vital artificial intelligence tools at their disposal without necessarily having to work with computer scientists.

7. Conclusions

The rendering of a bibliometric map, although computationally rigorous, is strongly affected by the researcher's judgment regarding the grouping of the Bibliometric Object (BO), the selection of the objective function, and the optimization algorithm to solve the problem. In this study, we introduced a novel framework for bibliometric analysis, accompanied by critical insights of the most influential papers. Particularly, this review paper presented the main outcomes of a bibliometric analysis for Adaptive Learning Systems based on 5564 research papers. We discussed the inter-associations of the BOs in the corresponding keywords' map, along with a narrative analysis of the most cited papers in (a) the entire database and (b) in the last five years. Furthermore, by using time-series analysis, we identified the trends of the top five keywords, the significant associations among them, and emerging trends. By combining this computational procedure and the conceptual analysis, we obtained a documented overview of the literature regarding Adaptive Learning Systems. A limitation of the research work is that although a large number of articles has been used for the generation of the papers' collection, these papers have been extracted from a single database only. Particularly, we used the Scopus service, which is considered to be one of the major scientific literature databases. Other databases could have been employed for the same, either on their own or in combination with the one used in the study, for example Web of Science, which is a rich and reliable alternative. The idea of enriching the study with additional data taken from other databases is a clear future research direction. In addition, it would be interesting to analyze each one of the specific subcategories identified in Table 2 using bibliometric algorithms in combination with critical analysis.

The vast amount of scientific papers that have been written and are constantly being produced confirm the interest of the academic community in developing more "smart" and "adaptive" digital learning systems. By combining the conclusions obtained from the bibliometric map and the time series of the appearance of the keywords, it was revealed that the research interest shifts from intelligent tutoring systems to adaptive learning systems through the utilization of the emerging technologies of artificial intelligence, learning analytics, and learning styles.

Author Contributions: Conceptualization, D.K., K.K., N.P.B., V.P., A.L. and S.A.C.; Methodology, D.K., V.P. and A.L.; Project administration, D.K. and K.K.; Resources, K.K.; Software, N.P.B.; Visualization, N.P.B.; Writing—original draft, N.P.B., V.P. and S.A.C.; Writing—review & editing, K.K. and A.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author, N.P.B., upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. The Most Frequent Keywords

Table A1 presents the 100 most frequent keywords found in the studied database, and their correction, in order of appearance grouped in the bibliometric map and time-series. The NULL identifier groups generic keywords, which were excluded from the analysis.

Table A1. 100 Most frequent keywords and their grouping.

Initial String	Corrected String	Frequency	Initial String	Corrected String	Frequency
adaptive learning	adaptive learning	1697	feedback	NULL	83
intelligent tutoring systems	intelligent tutoring system	1518	web 2.0	NULL	83
intelligent tutoring system	intelligent tutoring system	865	adaptivity	adaptive learning	81
personalized learning	personalized learning	673	distance learning	NULL	81
e-learning	NULL	665	knowledge tracing	knowledge management	81
machine learning	machine learning	282	learner model	learner model	80
learning	NULL	279	genetic algorithm	genetic algorithm	78
learning analytics	learning analytics	241	intelligent tutoring	intelligent tutoring system	78
neural network	artificial neural networks	240	adaptive learning systems	adaptive learning	75
artificial intelligence	artificial intelligence	236	learning management system	learning management system	71
personalization	personalized learning	232	learning objects	learning objects	71
neural networks	artificial neural networks	217	lifelong learning	NULL	71
ontology	ontologies	208	smart learning	intelligent tutoring system	70
personal learning environment	personalized learning	206	smart learning environments	intelligent tutoring system	70
mobile learning	NULL	193	affect	NULL	69
personal learning environments	personalized learning	171	artificial neural networks	artificial neural networks	69
online learning	NULL	169	problem solving	NULL	68
adaptive learning rate	adaptive learning	167	serious games	NULL	68
deep learning	artificial neural networks	167	adaptive	adaptive learning	67
education	NULL	163	e-learning	NULL	67
self-regulated learning	self-regulated learning	163	instructional design	NULL	66
reinforcement learning	reinforcement learning	158	ontologies	ontologies	65
collaborative learning	NULL	155	recommender systems	recommender systems	65
learning style	learning style	148	virtual reality	NULL	65
data mining	data mining	145	game-based learning	NULL	64
higher education	NULL	145	metacognition	metacognition	64
educational data mining	data mining	141	mooc	NULL	64
learning styles	learning style	141	Bayesian networks	Bayesian networks	61
student modeling	student modeling	132	adaptive hypermedia	adaptive learning	60
adaptation	adaptive learning	130	item response theory	item response theory	59
its	intelligent tutoring system	117	knowledge representation	knowledge management	59
artificial neural network	artificial neural networks	108	clustering	clustering	58

Table A1. *Cont.*

Initial String	Corrected String	Frequency	Initial String	Corrected String	Frequency
interactive learning environments	intelligent tutoring system	106	user modeling	user modeling	58
semantic web	NULL	106	cloud computing	NULL	57
adaptive control	adaptive learning	105	knowledge management	knowledge management	57
motivation	motivation	102	personalised learning	personalized learning	57
adaptive learning system	adaptive learning	100	bp neural network	artificial neural networks	55
student model	student modeling	100	simulation	NULL	55
affective computing	NULL	99	lms	learning management system	54
natural language processing	natural language processing	96	ubiquitous learning	NULL	54
evaluation	NULL	95	collaboration	NULL	53
assessment	NULL	94	moocs	NULL	53
blended learning	NULL	93	concept drift	concept drift	52
big data	data mining	89	intelligent tutoring system (its)	intelligent tutoring system	52
ple	personalized learning	89	m-learning	NULL	52
gamification	NULL	86	prediction	NULL	52
smart learning environment	intelligent tutoring system	85	augmented reality	NULL	51
classification	classification	84	adaptive e-learning	adaptive learning	50
fuzzy logic	NULL	84	collaborative filtering	recommender systems	50
educational technology	NULL	83	engagement	NULL	50

References

1. Koutsantonis, D.; Panayiotopoulos, J.C. Expert system personalized knowledge retrieval. *Oper. Res.* **2011**, *11*, 215–227. [CrossRef]
2. Mastorodimos, D.; Chatzichristofis, S.A. Studying Affective Tutoring Systems for Mathematical Concepts. *J. Educ. Technol. Syst.* **2019**, *48*, 14–50. [CrossRef]
3. Dima, A.; Bugheanu, A.M.; Dinulescu, R.; Potcovaru, A.M.; Stefanescu, C.A.; Marin, I. Exploring the Research Regarding Frugal Innovation and Business Sustainability through Bibliometric Analysis. *Sustainability* **2022**, *14*, 1326. [CrossRef]
4. Popescu, D.V.; Dima, A.; Radu, E.; Dobrotă, E.M.; Dumitrache, V.M. Bibliometric Analysis of the Green Deal Policies in the Food Chain. *Amfiteatru Econ.* **2022**, *24*, 410–428. [CrossRef]
5. Plevris, V.; Solorzano, G.; Bakas, N. Literature review of historical masonry structures with machine learning. In Proceedings of the 7th ECCOMAS Thematic Conference on Computational Methods in Structural Dynamics and Earthquake Engineering, Crete, Greece, 24–26 June 2019; ECCOMAS: Crete, Greece, 2019; pp. 1547–1562.
6. Plevris, V.; Bakas, N.; Markeset, G.; Bellos, J. Literature review of masonry structures under earthquake excitation utilizing machine learning algorithms. In Proceedings of the 6th ECCOMAS Thematic Conference on Computational Methods in Structural Dynamics and Earthquake Engineering, Rhodes Island, Greece, 15–17 June 2017; ECCOMAS: Rhodes Island, Greece, 2017; pp. 2685–2694.
7. Papadaki, M.; Bakas, N.; Ochieng, E.; Karamitsos, I.; Kirkham, R. Big data from social media and scientific literature databases reveals relationships among risk management, project management and project success. *PM World J.* **2019**, *8*. [CrossRef]
8. Dimopoulos, T.; Bakas, N. An artificial intelligence algorithm analyzing 30 years of research in mass appraisals. *RELAND Int. J. Real Estate Land Plan.* **2019**, *2*, 10–27.
9. Kruskal, J.B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* **1964**, *29*, 1–27. [CrossRef]
10. Moré, J.J. The Levenberg-Marquardt algorithm: Implementation and theory. In *Numerical Analysis*; Springer: Berlin, Germany, 1978; pp. 105–116.

11. Shepard, R.N. The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika* **1962**, *27*, 125–140. [CrossRef]
12. van Eck, N.; Waltman, L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* **2009**, *84*, 523–538. [CrossRef]
13. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
14. Conn, A.R.; Gould, N.I.; Toint, P. A globally convergent augmented Lagrangian algorithm for optimization with general constraints and simple bounds. *SIAM J. Numer. Anal.* **1991**, *28*, 545–572. [CrossRef]
15. Audet, C.; Dennis, J.E., Jr. Analysis of generalized pattern searches. *SIAM J. Optim.* **2002**, *13*, 889–903. [CrossRef]
16. Shi, Y.; Eberhart, R. A modified particle swarm optimizer. In Proceedings of the 1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence, Anchorage, AK, USA, 4–9 May 1998; IEEE: Piscataway, NJ, USA, 1998; pp. 69–73.
17. Plevris, V. Innovative Computational Techniques for the Optimum Structural Design Considering Uncertainties. Ph.D. Thesis, National Technical University of Athens, Athens, Greece, 2009.
18. Moayyeri, N.; Gharehbaghi, S.; Plevris, V. Cost-Based Optimum Design of Reinforced Concrete Retaining Walls Considering Different Methods of Bearing Capacity Computation. *Mathematics* **2019**, *7*, 1232. [CrossRef]
19. Byrd, R.H.; Gilbert, J.C.; Nocedal, J. A trust region method based on interior point techniques for nonlinear programming. *Math. Program.* **2000**, *89*, 149–185. [CrossRef]
20. Ugray, Z.; Lasdon, L.; Plummer, J.; Glover, F.; Kelly, J.; Martí, R. Scatter search and local NLP solvers: A multistart framework for global optimization. *INFORMS J. Comput.* **2007**, *19*, 328–340. [CrossRef]
21. Levenberg, K. A method for the solution of certain non-linear problems in least squares. *Q. Appl. Math.* **1944**, *2*, 164–168. [CrossRef]
22. Marquardt, D.W. An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* **1963**, *11*, 431–441. [CrossRef]
23. Bakas, N.P.; Plevris, V.; Langousis, A.; Chatzichristofis, S.A. ITSO: A novel inverse transform sampling-based optimization algorithm for stochastic search. *Stoch. Environ. Res. Risk Assess.* **2022**, *36*, 67–76. [CrossRef]
24. Plevris, V.; Bakas, N.P.; Solorzano, G. Pure Random Orthogonal Search (PROS): A Plain and Elegant Parameterless Algorithm for Global Optimization. *Appl. Sci.* **2021**, *11*, 5053. [CrossRef]
25. Pranckutė, R. Web of Science (WoS) and Scopus: The titans of bibliographic information in today's academic world. *Publications* **2021**, *9*, 12. [CrossRef]
26. Scopus Database. Available online: https://www.scopus.com/search/ (accessed on 17 March 2022).
27. Liu, W.H.; Sherman, A.H. Comparative analysis of the Cuthill–McKee and the reverse Cuthill–McKee ordering algorithms for sparse matrices. *SIAM J. Numer. Anal.* **1976**, *13*, 198–213. [CrossRef]
28. Gates, R.L. CuthillMcKee.jl. Available online: https://github.com/rleegates/CuthillMcKee.jl (accessed on 17 March 2022).
29. Anderson, J.R.; Corbett, A.T.; Koedinger, K.R.; Pelletier, R. Cognitive tutors: Lessons learned. *J. Learn. Sci.* **1995**, *4*, 167–207. [CrossRef]
30. VanLehn, K. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educ. Psychol.* **2011**, *46*, 197–221. [CrossRef]
31. Felder, R.M.; Silverman, L.K. Learning and teaching styles in engineering education. *Eng. Educ.* **1988**, *78*, 674–681.
32. Murray, T. Authoring intelligent tutoring systems: An analysis of the state of the art. *Int. J. Artif. Intell. Educ. (IJAIED)* **1999**, *10*, 98–129.
33. Koedinger, K.R.; Anderson, J.R.; Hadley, W.H.; Mark, M.A. Intelligent tutoring goes to school in the big city. *Int. J. Artif. Intell. Educ.* **1997**, *8*, 30–43.
34. Bloom, B.S. The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educ. Res.* **1984**, *13*, 4–16. [CrossRef]
35. VanLehn, K. The behavior of tutoring systems. *Int. J. Artif. Intell. Educ.* **2006**, *16*, 227–265.
36. Corbett, A.T.; Anderson, J.R. Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Model. -User-Adapt. Interact.* **1994**, *4*, 253–278. [CrossRef]
37. Brusilovsky, P. Methods and Techniques of Adaptive Hypermedia. In *Learner Modeling and Learner-Adapted Interaction*; Springer, Dordrecht, The Netherlands, 1996; Volume 6.
38. Hattie, J.; Timperley, H. The power of feedback. *Rev. Educ. Res.* **2007**, *77*, 81–112. [CrossRef]
39. Shute, V.J. Focus on formative feedback. *Rev. Educ. Res.* **2008**, *78*, 153–189. [CrossRef]
40. Aleven, V.; Mclaren, B.M.; Sewall, J.; Koedinger, K.R. A new paradigm for intelligent tutoring systems: Example-tracing tutors. *Int. J. Artif. Intell. Educ.* **2009**, *19*, 105–154.
41. Conati, C.; Maclaren, H. Empirically building and evaluating a probabilistic model of user affect. *User Model. -User-Adapt. Interact.* **2009**, *19*, 267–303. [CrossRef]
42. Piech, C.; Bassen, J.; Huang, J.; Ganguli, S.; Sahami, M.; Guibas, L.J.; Sohl-Dickstein, J. Deep knowledge tracing. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canads, 7–12 December 2015; Volume 28.
43. Strayer, J.F. How learning in an inverted classroom influences cooperation, innovation and task orientation. *Learn. Environ. Res.* **2012**, *15*, 171–193. [CrossRef]

44. Baker, R.S.; D'Mello, S.K.; Rodrigo, M.M.T.; Graesser, A.C. Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. *Int. J. Hum.-Comput. Stud.* **2010**, *68*, 223–241. [CrossRef]
45. VanLehn, K.; Lynch, C.; Schulze, K.; Shapiro, J.A.; Shelby, R.; Taylor, L.; Treacy, D.; Weinstein, A.; Wintersgill, M. The Andes physics tutoring system: Lessons learned. *Int. J. Artif. Intell. Educ.* **2005**, *15*, 147–204.
46. Schulze, K.G.; Shelby, R.N.; Treacy, D.J.; Wintersgill, M.C.; Vanlehn, K.; Gertner, A. Andes: An intelligent tutor for classical physics. *J. Electron. Publ.* **2000**, *6*. [CrossRef]
47. Conati, C.; Gertner, A.; Vanlehn, K. Using Bayesian networks to manage uncertainty in student modeling. *User Model. User-Adapt. Interact.* **2002**, *12*, 371–417. [CrossRef]
48. Chen, C.M.; Lee, H.M.; Chen, Y.H. Personalized e-learning system using item response theory. *Comput. Educ.* **2005**, *44*, 237–255. [CrossRef]
49. Papanikolaou, K.A.; Grigoriadou, M.; Kornilakis, H.; Magoulas, G.D. Personalizing the Interaction in a Web-based Educational Hypermedia System: The case of INSPIRE. *User Model. User-Adapt. Interact.* **2003**, *13*, 213–267. [CrossRef]
50. Brusilovsky, P.; Schwarz, E.; Weber, G. ELM-ART: An intelligent tutoring system on World Wide Web. In Proceedings of the International Conference on Intelligent Tutoring Systems, Montréal, Canada, 12–14 June 1996, Springer: Berlin, Germany, 1996; pp. 261–269.
51. Chen, C.M.; Chung, C.J. Personalized mobile English vocabulary learning system based on item response theory and learning memory cycle. *Comput. Educ.* **2008**, *51*, 624–645. [CrossRef]
52. Vrugt, J.A.; Robinson, B.A.; Hyman, J.M. Self-adaptive multimethod search for global optimization in real-parameter spaces. *IEEE Trans. Evol. Comput.* **2008**, *13*, 243–259. [CrossRef]
53. Castro, J.R.; Castillo, O.; Melin, P.; Rodríguez-Díaz, A. A hybrid learning algorithm for a class of interval type-2 fuzzy neural networks. *Inf. Sci.* **2009**, *179*, 2175–2193. [CrossRef]
54. Morze, N.; Varchenko-Trotsenko, L.; Terletska, T.; Smyrnova-Trybulska, E. Implementation of adaptive learning at higher education institutions by means of Moodle LMS. *J. Physics Conf. Ser.* **2021**, *1840-1* , 012062. [CrossRef]
55. Ullah, N.; Mugahed Al-Rahmi, W.; Alzahrani, A.I.; Alfarraj, O.; Alblehai, F.M. Blockchain technology adoption in smart learning environments. *Sustainability* **2021**, *13*, 1801. [CrossRef]
56. Molenaar, I.; Horvers, A.; Baker, R.S. What can moment-by-moment learning curves tell about students' self-regulated learning? *Learn. Instr.* **2021**, *72*, 101206. [CrossRef]
57. Han, J.; Kim, K.H.; Rhee, W.; Cho, Y.H. Learning analytics dashboards for adaptive support in face-to-face collaborative argumentation. *Comput. Educ.* **2021**, *163*, 104041. [CrossRef]
58. Alamri, H.A.; Watson, S.; Watson, W. Learning technology models that support personalization within blended learning environments in higher education. *TechTrends* **2021**, *65*, 62–78. [CrossRef]
59. Schiff, D. Out of the laboratory and into the classroom: The future of artificial intelligence in education. *AI Soc.* **2021**, *36*, 331–348. [CrossRef]
60. Hutt, S.; Krasich, K.R.; Brockmole, J.; D'Mello, S.K. Breaking out of the lab: Mitigating mind wandering with gaze-based attention-aware technology in classrooms. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 8–13 May 2021; pp. 1–14.
61. Tabuenca, B.; Serrano-Iglesias, S.; Carruana-Martin, A.; Villa-Torrano, C.; Dimitriadis, Y.A.; Asensio-Perez, J.I.; Alario-Hoyos, C.; Gomez-Sanchez, E.; Bote-Lorenzo, M.L.; Martinez-Mones, A.; et al. Affordances and core functions of smart learning environments: A systematic literature review. *IEEE Trans. Learn. Technol.* **2021**, *14*, 129–145. [CrossRef]
62. Christopoulos, A.; Sprangers, P. Integration of educational technology during the Covid-19 pandemic: An analysis of teacher and student receptions. *Cogent Educ.* **2021**, *8*, 1964690. [CrossRef]

economies

Article

The Influence of E-HRM on Modernizing the Role of HRM Context

Adambarage Chamaru De Alwis [1], Berislav Andrlić [2,*] and Marko Šostar [2,*]

[1] Faculty of Commerce and Management and Studies, University of Kelaniya, Kelaniya 11300, Sri Lanka; chamaru@kln.ac.lk
[2] Social Department, Polytechnic in Pozega, Vukovarska 17, 34000 Pozega, Croatia
* Correspondence: bandrlic@vup.hr (B.A.); msostar@vup.hr (M.Š.); Tel.: +385-995226371 (B.A.); +385-98554442 (M.Š.)

Abstract: Human Resource Management (HRM) practitioners must play multiple roles to achieve both functional and organizational objectives. In the current business environment, Electronic Human Resource Management (E-HRM) is a pervasive application of Information Technology (IT), and its global adoption is now widespread. This study's primary objective is to determine the effect of E-HRM implementation in Sri Lanka on the evolving HRM function. The study employed a qualitative methodology. On the basis of prior research, a study framework was formulated. After collecting the findings, the researcher analyzed the data using the study framework. To answer the research question and attain the research objectives, the researcher compared the theory to the findings during the analysis. The researcher determined that HRM was unfamiliar with E-HRM applications but was utilizing E-HR tools in their daily operations. According to the study, HRM places a greater emphasis on administrative tasks. The greatest benefit of implementing E-HRM practice is that it liberates HRM from intermediary roles, allowing them to focus on strategic planning in HR organizations and transforming HR practitioners from administrative paper handlers to strategic planners.

Keywords: electronic HRM (E-HRM); HR practitioners; administrative expert; strategic partner

Citation: De Alwis, Adambarage Chamaru, Berislav Andrlić, and Marko Šostar. 2022. The Influence of E-HRM on Modernizing the Role of HRM Context. *Economies* 10: 181. https://doi.org/10.3390/economies10080181

Academic Editors: Jian-Hong Ye, Yung-Wei Hao, Yu-Feng Wu and Savvas A. Chatzichristofis

Received: 29 June 2022
Accepted: 22 July 2022
Published: 27 July 2022

1. Background of the Study

Individuals are what make an organization function. As a result, it is necessary to recruit, hire, and retain potential employees (Laumer et al. 2010). Human Resource Management (HRM) departments manage all personnel-related issues affecting an organization's workforce, including employee performance, training, recruitment, staffing, and payments (Stolt 2010). Stolt (2010) defines HRM as a method for strategically managing the organization's most valuable asset, the individual who contributes to the business's success individually or collectively. HRM has historically been critical to an organization's performance improvement. According to Opatha (1995), HRM is the efficient and effective use of HRs to accomplish an organization's goals. There is a strong correlation between organizational effectiveness and HRM.

The evolution of technology over the past few decades has resulted in creative commercial developments. In addition, technology has altered how HR practitioners communicate, work, and manage employees. It has two direct effects on businesses: (1) an increase in efficiency, effectiveness, and productivity, and (2) a shift in the way people create, organize, manage, and operate a business (Zhang and Wang 2006). These significant technological developments and modernizations, together with the expansion of the internet and other specialized equipment, have compelled organizations to maximize their potential for performing and enhancing the performance and efficiency of their HRM departments. Consequently, Human Resource (HR) practitioners must adapt to the growing competition, changing employee attitudes, and rapid advancements in HR technology.

Since the invention of the internet, a new era of HRM known as Electronic Human Resource Management (E-HRM) has begun, radically reengineering HR practices and strategies in order to compete in an intensely competitive market. Strohmeier (2007) defines E-HRM as the "use of information technology to connect and support at least two individual or collective actors in performing HRM activities collaboratively". With the aid of E-HRM, it is anticipated that HR practitioners would be able to work more strategically and efficiently. Under this scenario, technology enables HR departments to address the organization's HR requirements through web-based channels. It allows the employees and HR practitioners to monitor, view, and make the required adjustments for more efficient and effective HR management. In addition, the implementation of E-HRM reduces the demand for HR specialists due to the fact that E-HRM eliminates the "HR middleman" (Lengnick-Hall and Moritz 2003). Now, firms can accomplish their duties with fewer staff.

HR practitioners interact with employees and participate in human resource activities to increase the effectiveness of their organization's provision of technical guidance and other human capital services. Traditionally, HR practitioners have been required to physically accomplish an abundance of administrative and paperwork tasks. In brief, this is ineffective and inaccurate comparatively. According to Beadles et al. (2005), E-HRM enables top management and HR practitioners to act as strategic partners as opposed to administrative specialists. Consequently, the role of HR practitioners has evolved alongside technological developments.

According to Ha (2011), the use of information systems in HRM is not novel; however, the rapid evolution of such technology inspired HRM. Consequently, the use of technology in HRM has increased in popularity. Kemske (2008) confirms that the pressures and priorities placed on HR departments are responsible for the abandonment of traditional practices and the emergence of organizational innovation.

The contemporary HR professional is anticipated to be more strategic and adaptable and economically and customer-focused (Snell and Lepak 2002). E-HRM initially focuses on the recruitment, selection, compensation, and evaluation processes even though it currently aims to implement all operational HRM activities. E-HRM initially focuses on the recruitment, selection, compensation, and evaluation processes even though it currently aims to implement all operational HRM activities. However, some scholars claim that even though the new E-HRM brings a significant challenge for the HR profession, in some instances, E-HRM capabilities have some similarities with traditional HRM duties, such as organizing organizations and jobs for people, obtaining HRs, enhancing employee motivation, and maintaining HRs (Fisher et al. 1996).

According to Ruel et al. (2007), a high level of commitment to moving HRM tasks, particularly administrative ones, to the electronic version has a significant chance of shifting the focus to one that is more strategically oriented. HR consultants have reported that both the number of companies implementing E-HRM and the scope of applications within organizations are consistently rising. Businesses are just starting to use information technology (IT) to deliver superior HRM services (De Alwis 2010). The adoption of E-HR changed HR's focus from operational HR to being a more strategic partner to the organization by becoming more involved in strategic decision-making, where before, HR practitioners served as administrative experts and employee champions (Bell et al. 2006). The importance of technological advancements in shifting the HR department's role from administrative to strategic has been emphasized in numerous studies worldwide.

According to De Alwis (2010), 70% of companies have moderate knowledge and usage of E-HR at different levels. However, some of them are still in the beginning stages by only using a small number of their functions, such as managing leave. This study attempts to provide critical points that, hopefully, will help to clarify the role of E-HRM in HRM in the Sri Lankan context. Again Marler and Fisher (2013) conducted a study titled "Influence of E-HRM on Renovating HRM's Role" and disclosed that they were unable to find empirical evidence that E-HRM predicts strategic outcomes in Sri Lanka. In 2017, Mdhushani and De

Alwis (2017) reaffirmed their earlier findings. HR pioneers and local findings are in conflict with one another.

Chandradasa and Priyashantha (2021) found that employees are not overly concerned with the compliance level and satisfaction received by utilizing E-HRM when deciding whether to use E-HRM for their work. Obeidat (2016) conducted a study titled "The link between E-HRM use and HRM effectiveness: an empirical study" and discovered that the use of E-HRM has a positive impact on HRM effectiveness at both the policy and practice levels. Furthermore, it confirms the role of user intention in mediating the relationship between E-HRM determinants (both performance expectancy and social influence) and E-HRM use. Thathsara and Sutha (2021) discovered that E-HRM practices have a substantial and favorable effect on organizational performance and that organizational agility mediates the link between E-HRM practices and operational performance. Mdhushani and De Alwis (2017) concluded that managers have some understanding of HR technology, but many are unfamiliar with the phrase E-HRM. It is projected that they would automate more HR processes and implement modern HR technologies in an effort to obtain more significant benefits. According to the conclusions of Kumara and Galhena's research, the major factor influencing HRIS program utilization is top management support (Kumara and Galhena 2021). In addition, it was determined that perceived utility, IT skill, and subjective standards are key HRIS utilization factors. Another study on the efficacy of E-HRM systems in Sri Lanka was undertaken by Weerasuriya (2006), and the results revealed that the type of these systems and their utilization levels or frequency varied from one company to another. Additionally, it demonstrated that there was a difference between actual and ideal deployment levels, suggesting that existing systems or modules could still benefit from additional enhancements. The researcher identified the work systems that are significantly impacted by the deployment of E-HRM. De Alwis et al. (2019) discovered that internal environmental factors were once again substantially correlated with the amount of HRIS application adoption in 2019. This, in turn, has an effect on the level of application adoption for HRIS. There are different varieties of E-HRM-related research studies. However, no study has examined how E-HRM influences the evolution of HR professionals' roles, especially in a Sri Lankan Context.

2. Problem Statement

This study is aimed at how E-HRM influences HR practitioners' roles. Because of technological changes, HR practitioners had to learn to deal with changing attitudes in the workforce and become more flexible and cost-effective. Thus, when E-HRM is implemented and developed, it is critical to monitor both positive and negative effects, such as changes in the roles and capabilities of HR practitioners. As such, the research will concentrate on the following issues: What are the HR functions that E-HRM has primarily served, and how many HR practitioners have transformed as E-HRM has been implemented in Sri Lanka?

3. Significance of the Study

In a growing number of organizations, electronic HRM is replacing face-to-face HR duties, albeit at varying levels. Additionally, E-HRM enables the HR function to be dynamic and efficient. In this context, a comprehensive examination of the evolution of technology in human resources and its implications is required.

Initially, it is essential to determine the significance of the study for organizational effectiveness. E-HRM enhances the internal profile of the HR department, thereby improving the work environment. E-HRM improves system visibility. It increases the efficacy and cost-effectiveness of human resources and positions HR as a strategic partner in achieving organizational objectives. It is obvious that information technology is currently transforming the HR department, especially in terms of how HR services are provided. Even medium-sized businesses now have internet access, which is used to provide HR services. E-HRM appears to be integrated into every aspect of HRM, including planning, administration, and communication.

Studies on the use of E-HRM typically concentrate on American and European examples. Despite the complexity and cultural shifts, they exhibit some consistency. The majority of developed countries, however, are spread out across distinct cultural regions. Jordan's E-HRM is still in its early stages, so the country's IT environment is evolving (Obeidat 2016). This study will assist in understanding HR professionals' current role in developing nations regarding transfer-HRM. It will contribute to the expansion of knowledge in the literature.

The primary benefit of implementing E-HRM is that it liberates HR practitioners from intermediary roles, allowing them to focus on strategic planning within their HR organization and the changes brought about by HR practitioners' transition from administrative paper handlers to strategic planners. This study can help HR professionals by improving their understanding of E-HRM practices. Furthermore, they can be more confident in implementing E-HRM practices. By increasing knowledge about minimizing the risk associated with E-HRM and demonstrating the benefits of E-HRM, the study benefited a variety of parties, including the organization's top management, HR practitioners, and software vendors.

Finally, it aided in understanding the history of E-HRM and its relationship to the role of HR practitioners.

4. Literature Review

There is little doubt that "people" are a critical asset that determines an organization's success or failure, and thus the importance of those people's knowledge, skills, attitudes, and behaviors are very important for the organization's advancement (De Alwis 2010). According to Gebauer (2003), the company's staff, which is equipped with all of the necessary skills and knowledge, has become a critical factor in its success. According to Buckley et al. (2004) and Mdhushani and De Alwis (2017), effective management of human talent and intellectual capital will be a critical factor in determining shareholder value. According to Attwood (1989), personal management is the aspect of management concerned with the management of people at work. Human Resource Management (HRM) is the process of attracting, selecting, retaining, developing, and utilizing human capital to achieve personal and organizational goals (Cascio and Awad 1981). Human Resource Management (HRM) is defined by Fisher et al. (2006) as "all management decisions and practices that have a direct impact on or influence the people, or HRs, who work for the organization".

In any organization, HR professionals play a critical role. The more roles it fulfills, the more likely that it will be effective in increasing the organization's productivity, enhancing the organization's quality of life at work, and adhering to all applicable laws and regulations governing HR utilization (Kramar et al. 1998). According to Opatha (2009), traditionally, HR practitioners were responsible for developing staffing plans, providing specific job training, managing annual performance appraisal programs, and handling day-to-day administrative tasks. According to Fisher et al., many firms retain a traditional personal department, which is frequently physically and psychologically isolated from the organization's "real work", and personal activities and staff are relatively isolated from the organization's "profit-making heart".

Numerous academics and practitioners have identified two primary roles for HR professionals: the administrative and traditional role and the strategic role (Beer 1997; Truss 2008). According to Fisher et al. (2006), HR professionals must be involved in identifying better ways to share services and reengineer administrative and other processes within the firm and across the organization.

E-HRM is a method of implementing HR techniques, policies, and practices in organizations by utilizing and/or maximizing the use of web-based technology-based channels (Ruel et al. 2007). E-HRM technology creates a portal that enables practitioners, employees, and HR practitioners to view, extract, and modify data required for managing the organization's HRs. Strohmeier (2007) defines E-HRM as the use of information technology to facilitate networking and to assist at least two individual or collective actors in performing HRM activities collaboratively. According to surveys, the number of organizations adopting E-HRM is increasing, as is the breadth of applications within those organizations.

"E-HRM has the potential to improve both the efficiency and effectiveness of the HRs function". Reducing office work cycle times, increasing data accuracy, and reducing the HR workforce can increase efficiency. Improving practitioners' and workers' ability to make better, more appropriate decisions can increase efficiency.

Furthermore, E-HRM enables HR to explore innovative ways to contribute to organizational effectiveness through knowledge management and the development of intellectual and social capital (Lengnick-Hall and Moritz 2003).

Based on the previous definitions, we can say that E-HRM is the use of information technology in HR practices to make it easier for employees and employers to communicate with one another and for organizations to improve their HR skills.

Since the inception of research on the relationship between web-based technologies and HRM, a variety of definitions have been proposed to explain the phenomenon dubbed E-HRM, which was previously referred to as an HR information system, virtual HRM, or web-based HRM (Fisher et al. 2006). The term was coined in response to the rise of e-commerce in the business world, and E-HR is highly specific to internet usage. Thus, Kettley and Reilly (2003) suggest that the more accurate term is "online HRM". Some users are initially perplexed by the terms HRIS and E-HRM. However, HRIS refers to systems used within the HRs department to enhance HR processes, whereas E-HR is intended to benefit non-HRs personnel such as employees and practitioners (Ruel et al. 2007). E-HR and E-HRM appear to be synonymous terms in some contexts. Academics frequently use the term "E-HRM", whereas practitioners, IT professionals, and software suppliers frequently use the terms "self-service" and "E-HR" (Panayotopoulou et al. 2007).

There is little consensus on how to define E-HRM, and the concept varies depending on the researcher's viewpoint (Strohmeier 2007; Bondarouk and Looise 2009). According to Kettley and Reilly (Panayotopoulou et al. 2007), E-HRM is "the use of traditional, web, and voice technologies to improve HR administration, transactions, and process performance". E-HRM is "the administrative support of the HR function in organizations through the use of Internet technology", according to Voermans and Van Veldhoven (2007). They also mention how the implementation of E-HRM might change how the HR function is carried out.

E-HRM is described as "the application of any technology that enables practitioners and employees to have direct access to HRs and other workplace services for communication, performance reporting, team management, knowledge management, and learning" (Opatha 1995) in addition to administrative applications. "Planning, implementing, and applying information technology for the purpose of networking and assisting at least two individual or collective actors in performing HR activities collaboratively", according to Strohmeier (2007), is what is meant by the term "E-HRM". A "completely integrated, enterprise-wide electronic network of HR-related data, information, services, databases, tools, applications, and transactions", according to Foster (2009), is what the phrase refers to.

E-HRM should be approached more strategically, per some definitions. E-HRM is "an umbrella term that encompasses all possible mechanisms and content for integrating HRM and information technology with the objective of creating value for targeted employees and management within and across organizations", according to Bondarouk and Looise (2009). This definition will be used in this study because it covers both the broad strategic outcomes that are the focus of this investigation as well as the administrative elements of E-HRM.

E-HRM is both a theory and a practice in HRM (Ruel et al. 2007). E-HRM can be viewed as a synthesis of HRM and information technology, above and beyond any of these definitions. Employing technological tools or web-based technologies can help organizations improve the efficacy and efficiency of their HR practices and policies.

In recent years, an increasing number of HR practitioners have placed a premium on technological advancements in HR functions, most notably in the reorganization of HR roles and responsibilities. Evidence of this shift has been discovered in studies (De Alwis 2010), which confirmed that HR practitioners' roles will significantly change as strategic partners in business. As per De Alwis (2010), HR professionals serve as strategic partners (50%), employee champions (30%), and administrative experts (20%).

By reducing administrative burdens, HR technology enables a shift in the emphasis of HR work from transactional to strategic. This is the main impact of HR technology on HR roles. Technology is seen by HR professionals as a way to advance their careers and take on a more strategic role within their organizations. As E-HRM eliminates the HR middleman, the switch from traditional HRM to E-HRM may also point to a reduced need for HR practitioners (Lengnick-Hall and Moritz 2003). It is also evident that before the adoption of E-HRM, HR professionals served as administrative experts and employee champions, and that after the adoption of E-HRM, HR's focus shifted away from operational HR and toward being a strategic partner to the organization by participating more in strategic decision making (De Alwis 2010). Retained HR professionals can assume more strategic responsibilities as a result of this transformation, which can be seen as an opportunity.

This shows the general scenario. The small number of earlier researchers who conducted their research in the Sri Lankan context, however, discovered conflicting findings. They assert that regardless of how E-HRM is applied, there is no proof that the role shifts from an administrative expert to one with a strategy-focused mindset.

For instance, the author cited a finding from a survey of web self-service deployment (Marler 2009) that companies have tended to use the web as a tactical tool to deliver HR services rather than as a means of rethinking fundamental operations and strategy and achieving more intangible benefits that come with the large and sustained investment. His argument about the possibility of E-HRM transforming HR into a strategic function is, in reality, a hard sell to the majority of practitioners who remain unaware of the potential and outcomes of E-HRM. Parry and Tyson (2011) recently reported on additional E-HRM outcomes, examining whether organizations actually accomplished the transformational goal of E-HRM. They argued that while there was no subjective evidence of a shift in HR's strategic role, the introduction of E-HRM did facilitate a shift in HR focus in terms of time spent on administrative versus strategic tasks in businesses.

To summarize, academics and practitioners disagree on whether HR practitioners have evolved into strategic partners as a result of E-HRM. While some researchers believed in the ability of E-HRM to elevate the HR role, others argued that E-HRM had not been recognized as a strategic decision maker in the HR department.

5. Methodology

This was a cross-sectional field study in which the researchers exerted little control over the sample. The inductive approach was used in this study because the researcher was concentrating on a single subject and attempting to reach a conclusion through analysis. Interviews were conducted to ascertain the public's perception of E-HRM practices because this is an exploratory study in which data was gathered using qualitative methods such as semi-structured interviews.

The primary study included organizations that had already implemented E-HRM functions. However, it did not specify whether they had completed the shift or were still in the process. Therefore, this study's population included all HR practitioners working in E-HRM-enabled organizations in Sri Lanka. The request is then forwarded along with our requirements. Additionally, we have made a strong request to designate one practitioner who is directly responsible for E-HRM operations. As a result, the study enrolled eight randomly chosen HR practitioners from companies who confirmed their involvement with E-HRM and were willing to participate in the subsequent discussion.

Due to ethical considerations, we will not recognize the organizations by name (they would be referred to as "company", "1", "2", etc.), and Table 1 exhibits their current position and number of years of experience.

Table 1. Profiles of the Interviewees.

No	Position	Gender	Company	Time Working at the Company
01	Assistant Manager—HRIS	Male	Company 1	2.5 years
02	Assistant Manager—HR	Female	Company 2	6.5 years
03	Assistant Manager—HR	Male	Company 3	2 years
04	Assistant Manager—HR	Female	Company 4	4.5 years
05	Manager HR and Administration	Male	Company 5	10 years
06	Senior Manager—HR	Male	Company 6	11 years
07	Manager HR and Administration	Male	Company 7	5 years
08	Manager—HR	Male	Company 8	4 years

Source: Research Data, 2022.

Multiple techniques are used to collect the data, a process called triangulation. Triangulation verifies that the data produced is accurate. Interviews and documentation reviews were conducted in this case. Documentation reviews are used to gain a thorough understanding of the current state of E-HRM adoption. The interview questions are based on Ha's survey (Ha 2011). All interviews took place over the phone, and the audio was captured using software. During the interviews, notes on significant responses were taken and incorporated into the transcripts.

The most efficient method for gathering data for this study is an interview because it is qualitative and inductive. We contacted HR practitioners who are experts in E-HRM in order to speed up the process of achieving goals and conducting research. This method of data collection had the advantage of allowing respondents to give thorough information about the subject. Each interview took place over a brief period of time. The questions can be rearranged in whatever order to suit the topic of the conversation. Clarifications and explanations were given to interviewees to help prevent misunderstandings.

The interviews were recorded and transcribed for further analysis. Although the data were analyzed using a process categorization technique. This procedure builds on the framework of the study, which is based on the literature. To begin, the data were organized according to the framework of the study. To gain a basic understanding of the data, the data analysis process began with a quantitative analysis of listings and categorizations.

It is critical for all researchers to be aware of ethical research practices. The term "ethics" refers to two distinct groups of people: those conducting research and those with fundamental rights that should be protected. As a result, this study had to be conducted fairly and impartially, avoiding all potential risks. Respondents must be informed of their legal rights. Respect for the individual is a fundamental human right. They have the option of participating or not participating in the research. Respondents were given the opportunity to act independently by providing pertinent information in order to participate in the study. Prior to that, the purpose of the study was thoroughly explained to them in the language they preferred. Confidentiality is a fundamental ethical principle, and anonymity is one way to maintain confidentiality. Anonymity was achieved in this study by omitting names from the study. The interviews took place in a private office and over the phone with their permission, and no third party was present during the conversation. Preventing harm is another fundamental human right to consider. To accomplish this, the researcher kept the interview time with the participants to a minimum. Keeping the interview private, confidential, and anonymous also prevented psychological harm. Participants in this study were treated fairly by providing them with information prior to their participation. Further, because the sample was chosen based on the guidelines for the study, everyone who met the criteria had a fair chance of being picked to take part in the study.

6. The Impact of E-HRM on the Roles of HR

The primary impact of HR technology on HR roles is that it allows HR work to shift from transactional to strategic by reducing administrative burdens. Human resource practitioners see technology as a source of revenue that allows them to play a more strategic role within their organizations. The interview questions were designed to elicit HR practitioners' perspectives on the benefits and drawbacks of E-HRM for the HR Manager role.

> "This technology enables us to manage our leaves efficiently, and employees can access their over-times via employee portals. We can save time, increase our effectiveness and efficiency, and reduce labour costs automatically. However, in the past, we were required to perform all tasks manually. It takes a long time, and occasionally the data is incorrect. Additionally, it lowers labour costs. For instance, we operate five businesses and utilize five payroll systems. As a result, we must employ five employees to monitor the system." (Assistant Manager, Human Resources-Company 1)

> " . . . I am no longer bound by monitor leaves. As a result, we save time. Prior to the system's implementation, we had to monitor leaves on a weekly basis . . . now it's simple to manage our administrative work." (Assistant Manager, Human Resources, Company 8)

> "It is extremely simple and quick for us. Additionally, it saves us time." (Director of Human Resources and Administration-Company 7)

According to the surveys of HR consultants, both the number of companies adopting E-HRM and the intensity of applications within those organizations continue to grow. IT is in the middle of assisting businesses in providing superior HR services. It is self-evident that information technology is transforming the HRs department at present, particularly in terms of the way HR services are delivered. Even medium-sized businesses now have access to the internet, which is used to provide HR services. E-HRM seems to be involved in every part of daily HRM, from planning and administration to communication.

All respondents agreed that their roles had been reshaped as a result of implementing the systems. E-HRM enables employees and HR practitioners to monitor, view, and make necessary adjustments to HRM. Additionally, with the use of E-HRM, fewer HRs professionals are required, as E-HRM eliminates the "HRs middleman" (Lengnick-Hall and Moritz 2003).

Ha (2011) asserts that today's Assistant Manager—HR must be more strategic, adaptable, cost-effective, and customer-oriented.

7. The Level of Influence That HR Technology Causes on the Role of HR

Much of the research emphasized the critical role of HR technology advancements in recasting the HR department's role from administrative to strategic. As a result of the adoption of E-HR, HR practitioners have taken on the roles of administrative expert and employee champion, and the adoption of E-HR has shifted HR's focus toward being a strategic partner of the organization by becoming involved in strategic decision-making (De Alwis 2010). The individual who fills this role should be capable of providing high-quality services at the lowest possible cost. Human resource practitioners view technology as a revenue stream that enables them to take on a more strategic role within their organizations. The shift from traditional HRM to E-HRM may also indicate that fewer HR practitioners are required, as E-HRM eliminates the HR middleman (Lengnick-Hall and Moritz 2003).

Ulrich's (1997) model outlines the various roles that HR departments play, including change agent, strategic partner, employee champion, and administrative expert. An administrative expert refers to the traditional HR role of designing and implementing HR processes such as recruiting, rewarding, training, hiring, and compensating. The majority of respondents agreed that implementing E-HRM would simplify administrative processes and improve HR service delivery. Since the HR professional will no longer have to worry about administrative tasks, he or she will be able to focus on the big picture.

As a result, I can devote my free time to strategic work. I am capable of supervising employees and making decisions regarding development, leadership development, and training. (Assistant manager-HRs-companies 8) As a result, I can devote my free time to strategic work. I am capable of supervising employees and making decisions regarding development, leadership development, and training. (Assistant Manager—HR HRs-companies 3) It's extremely simple for us, and it significantly reduces our manual workload and saves us time. That time can be used for analysis. (Assistant Manager—HR-Human Resources-Company 2) It alleviates some of our administrative burdens. (Assistant Manager—HR-Company 8).

The responses imply that HR practitioners are increasingly shifting their focus away from administrative expertise and toward strategic analysis. The strategic role is responsible for aligning HR strategies and practices with business strategies. This role identifies the organization's top talent, assists in filling job vacancies, communicates HR goals to employees to ensure they are implemented across the organization, and contributes to overall workplace productivity and harmony. Further, it puts strategies into action as quickly as possible by using the business's intended strategic direction and showing how HR's unique set of skills and competencies can help put strategies into action (Kirkbride 2003). Human resource practitioners emphasized the importance of utilizing the time saved for employee training programs and administrative tasks.

> "As a result, I can devote my free time to strategic work. I am capable of supervising employees and making development, leadership, and training decisions". (Assistant Manager, Human Resources, Company 5)

This finding indicates that HR's role as a strategic partner is evolving, though it supports HR's emphasis on the role of employee champion. Due to the reduction in workload, one respondent implies that he can gain new knowledge because of technological advancements. Paying special attention to the agent of change in the role of HR practitioners, E-HRM also lets the HR department investigate new ways of improving the effectiveness of the organization, such as knowledge management and building up intellectual and social capital (Lengnick-Hall and Moritz 2003).

> "We can acquire new knowledge about emerging technologies and generate new ideas about technological advancements". (Company 6 senior Manager—HR)

All respondents agreed that the implementation of the systems had reshaped their roles. Employees and HR practitioners can use E-HRM to monitor, view, and make changes to HRM.

A change agent is tasked with the responsibility of facilitating changes in other departments in order to maintain a company's competitive edge. This position communicates organizational changes to the rest of the organization. Ulrich (1997) advocated the use of change agents to assist with adjustment. They comprehend critical approaches to change, cultivate a commitment to those strategies, and ensure that proposed changes occur. These individuals coordinate training opportunities for employees to acquire new skills.

The organizational application of E-HRM practitioners must influence employee attitudes. As a result, they must organize awareness and training programs to ensure that the system is understood clearly. The Asst. manager—HRIS (Company 1) confirmed this notion by stating,

> "As a result of E-HRM adoption, we must train employees. As HR practitioners, we are confronted with a massive problem when it comes to training them. Therefore, we conduct awareness campaigns and educate employees about the system's benefits, gradually changing their mindset."

Change agents are concerned with the collective, whereas employee champions are concerned with individuals. A change agent essentially assists an organization in adapting to its next stage. Furthermore, prior to the implementation of the E-HRM, HR practitioners served as administrative experts and employee champions, and the implementation of the E-HRM shifted HR's focus from operational HR to be a strategic partner to the organization by involving HR in strategic decision-making (De Alwis 2010). This transformation can be viewed as an opportunity because retained HR practitioners will be able to take on more strategic roles. According to Ulrich (1997), the most desirable position for an HR manager is that of a "business partner", a term that is sometimes used instead.

However, the senior manager of HRs (Company 4) asserts that the implementation of technology has had no effect on the HR function. It has no effect on their changing responsibilities.

> "It does not allow us to improve the quality of our human resources department while not reducing our responsibilities. This allows me to obtain accurate data from the system as an HR manager. It speeds up my analysis and simplifies the report-creation process. As a result of technology implementation, our responsibilities have shifted. Because of the fluid nature of the HR manager's role".

Human resource practitioners appear to be taking a more proactive role and adding value to the organization by developing employee capabilities.

Technological advancements have changed how various HR functions are carried out. As a result, the typical HR job will be reshaped, shifting from administrative to strategy development. The most well-known business benefit of E-HR is that it allows HR professionals to refocus on their role as strategic business partners.

According to Edward and Mohrman (2003), when an organization has a fully integrated HRM system, HR is more likely to be a full partner in the strategic process. Having such a system, however, does not guarantee that HR will be a strategic partner, as some of the companies comparatively close to finalizing the integration confirm that now they have an opportunity to participate in decision-making.

Most respondents agreed that web-based tools for payroll, employee databases, and leave management were fully utilized. The majority of businesses in Sri Lanka are still in the process of upgrading their systems. The systems will be upgraded.

> "Indeed, we attempted to use an oracle system, but it failed due to its inability to capture our environment. We're currently attempting to install the h-senid system . . . ". (Assistant manager, Company 8)

> "Now, we are utilizing HRIS Micro Image. However, we intend to migrate it to a group-based system. Additionally, we utilize Micro Image's previous version as well as their new version, dubbed "Cloud Version System". (Assistant Manager—HR, Company 2)

More companies in Sri Lanka are incorporating electronic tools into their HRM processes. Additionally, businesses considering implementing these new forms of HRM should determine whether the results are superior to those obtained through previous HR processes. It is clear that the changes in technology in HRs and their impact should continue to grow. The main benefit of E-HRM is that it frees HR practitioners from intermediary roles and lets them focus solely on strategic planning in HR organizations. This changes HR practitioners from administrative paper handlers to strategic planners. The summarized findings are illustrated in Table 2.

Table 2. Summary of findings.

Dimensions	Main Findings
The impact of E-HRM on the roles of HR	• Low applications of technology • A transition from administrative expert to change agent and strategic partner

Source: Discussion Findings, 2022.

8. Conclusions

The majority of HR professionals' time is currently spent as administrative experts; however, the primary benefit of implementing E-HRM practices is that it frees HR professionals from intermediary roles, as a result of which the typical HR professional's responsibilities will shift from administrative to strategy development. It reaffirmed the findings of (Marler and Fisher 2013). In addition, businesses contemplating implementing these new forms of HRM should evaluate whether the outcomes outweigh those of previous HR procedures. The findings suggest that modifications to HR's role as a strategic partner may facilitate the shift in HR's focus to the role of employee champion. Online resources are rarely used for human resource activities such as learning and training. E-HR tools are mainly utilized in the recruitment and performance review processes.

Following are a few recommendations that the researcher would like to make in light of the analysis portion of the research's findings. E-HRM has an impact on how HR practitioners perform their jobs, so businesses should make more investments in it and aim to buy the most recent models. Organizations should concentrate increasingly on strategic HRM issues and on achieving strategic HR objectives by utilizing new technology and mastering administrative effectiveness. It is advised that senior management in the organizations make sure adequate funding is set aside for training activities to increase understanding of the term "E-HRM" and the significance of the E-HR tools.

This study sets the stage for future research in this area, which should focus on a variety of issues. First and foremost, more in-depth research with a larger sample size should be conducted to investigate the role of E-HRM in organizations. They also advise future researchers to supplement the study with quantitative methods.

Author Contributions: Conceptualization, A.C.D.A.; methodology, B.A. and M.Š.; formal analysis, A.C.D.A., B.A. and M.Š.; investigation, A.C.D.A., B.A. and M.Š.; resources, B.A. and M.Š.; data curation A.C.D.A.; writing—original draft preparation, B.A. and M.Š.; writing—review and editing, A.C.D.A.; supervision B.A.; project administration M.Š.; funding acquisition; B.A. and M.Š. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets presented in this study are not publically available due to ethical and privacy issues.

Conflicts of Interest: The authors declare no conflict of interest.

References

Attwood, Margaret. 1989. *Personnel Management*. London: Macmillan Education Limited.

Beadles, Nicholas Aston, II, Christopher M. Lowery, and Kim John. 2005. The impact of HR information systems: An exploratory study in the public sector. *Communications of the IIMS* 5: 6.

Beer, Michael. 1997. The transformation of the HR function: Resolving the tension between a traditional administrative and a new strategic role. *Human Resource Management* 36: 49–56. [CrossRef]

Bell, Bradford, Lee Sae-Won, and Yeung Sarah K. 2006. The impact of e-HR on professional competence in HRM: Implications for the development of HR practitioners. *Human Resource Management* 45: 295–308. [CrossRef]

Bondarouk, Tanya, and Jan Kees Looise. 2009. A Contingency Perspective on the Implementation of E-Performance Management. In *Encyclopedia of Human Resources Information Systems: Challenges in E-HRM*. Edited by Teresa Torres-Coronas and Mario Arias-Oliva. Hershey: IGI Global Books, pp. 97–202.

Buckley, Patrick, Kathleen Minette, Dennis Joy, and Jeff Michaels. 2004. The Use of Automated Employment Recruiting and Screening System for Temporary Professional Employees: A Case Study. *HRM Journal* 43: 233–41. [CrossRef]

Cascio, Wayne F., and Elias M. Awad. 1981. *HRM: An Information Systems Approach*. Reston: Florida International University.

Chandradasa, A. H. Isuru, and K. Gayan Priyashantha. 2021. Key Determinants of E-HRM Adoption Behavior: Banking Sector in Sri Lanka. Semantic Scholar. Retrieved 16 July 2022. Available online: https://www.semanticscholar.org/paper/Key-Determinants-of-E-HRM-Adoption-Behavior%3A-Sector-Chandradasa-Priyashantha/66d37478dfc6e7a67e013d200de34d3613f244b2 (accessed on 1 July 2022).

De Alwis, Adambarage Chamaru. 2010. The impact of electronic HRM on the role of HR practitioners. *Ekonomika A Management* 4: 47–59.

De Alwis, Adambarage Chamaru, Berislav Andrlić, and Marko Sostar. 2019. Internal Environment factors and the level of adoption of HRIS. Paper present at 30th DAAAM International Symposium on Intelligent Manufacturing and Automation, Zadar, Croatia, October 23–26.

Fisher, Cynthia D., Lyle F. Schoenfeldt, and J. B. Shaw. 2006. *Managing Human Resources*. Ner Delhi: Cengage Learning India (pvt) Ltd.

Fisher, Cynthia D., Schoenfeldt Lyle F., and J. B. Shaw. 1996. *An Introduction to Human Resource Management*. Boston: Houghton Mifflin Company.

Foster, Steve. 2009. Making Sense of e-HRM: Transformation, Technology and Power Relations. In *Handbook of Research on E-Transformation and Human Resources Management Technologies: Organizational Outcomes and Challenges*. Edited by Tanya Bondarouk, Huub Ruel, Karine Guiderdoni-Jourdain and Ewan Oiry. Information Science Reference. New York: Hershey, pp. 1–19.

Gebauer, Margaret. 2003. Information systems on human capital in service sector organizations. *New Library World* 104: 33–41. [CrossRef]

Ha, Ngo Thi Vo. 2011. The Impact of E-HRM on the Roles and Competencies of HR. Master's thesis, University of Vaasa, Vaasa, Finland.

Kemske, F. 2008. HR's role will change. The question is how? *Workforce* 77: 46–60.

Kettley, Polly, and P. Reilly. 2003. *An Introduction to e-HR*. Brighton: Institute of Employment Studies.

Kirkbride, Paul. 2003. Management development: In search of a new role? *Journal of Management Development* 22: 171–80. [CrossRef]

Kramar, Robin, Phil Mcgraw, and Schuler Randall. 1998. HRM in Australia. *Asia Pasific Journal of Human Resources* 36: 109–11.

Kumara, W. H. S., and Bandula L. Galhena. 2021. Determinants of Human Resource Informatıon System Usage: Evıdence From Ceylon Electrıcıty Board (CEB), Srı Lanka. *Kelaniya Journal of Human Resource Management* 16: 85. [CrossRef]

Laumer, Sven, Andreas Eckhardt, and Tim Weitzel. 2010. Electronic HRM in an ebusiness environment. *Journal of Electronic Commerce Research* 11: 240–50.

Edward, E. Lawlers, and Susan Alber Mohrman. 2003. *Creating a Strategic HR Organization: An Assessment of Trends and New Directions*. Stanford: Stanford.

Lengnick-Hall, Mark L., and Steve Moritz. 2003. The Impact of e-HR on the HRM Function. *Journal of Labour Research* 24: 365–79. [CrossRef]

Marler, Janet H. 2009. Making HRs strategic by going to the Net: Reality or myth? *The International Journal of HRM* 20: 515–27.

Marler, Janet H., and Sandra L. Fisher. 2013. An Evidence-Based Review of e-HRM and Strategic Human Resource Management. *Human Resource Management Review* 23: 8–36. [CrossRef]

Mdhushani, M. Kalani, and Adambarage Chamaru De Alwis. 2017. The Relationship between Internal Environmental Factors and the Level of Adoption of Human Resource Information System (HRIS) Applications in Manufacturing Sector Organizations in Colombo District, Sri Lanka. Paper present at 3rd HRM Student Research Symposium, Sarasavi Mawatha, Sri Lanka, July 25; pp. 34–35.

Obeidat, Shatha M. 2016. The link between e-HRM use and HRM effectiveness: An empirical study. *Personnel Review* 45: 1281–301. [CrossRef]

Opatha, Henarath H. D. N. P. 1995. *HRM Personnel*. Colombo: European Institute of Professional Education.

Opatha, Henarath H. D. N. P. 2009. *HRM: Personnel*. Colombo: Department of HRM, University of Jayewardenepura.

Panayotopoulou, Leda, Galanaki Eleanna, and Papalexandris Nancy. 2007. E-HR adoption and the role of HRM: Evidence from Greece. *Personnel Review* 36: 277–94. [CrossRef]

Parry, Emma, and Shaun Tyson. 2011. Desired goals and actual outcomes of E-HRM. *HRM Journal* 21: 335–54. [CrossRef]

Ruel, Huub J., Tanya Bondarouk, and Mandy Van der Velde. 2007. The contribution of e-HRM to HRM effectiveness: Results from a quantitative study in a Dutch Ministry. *Employee Relation* 29: 280–91. [CrossRef]

Snell, Scott A., and David P. Lepak. 2002. Virtual HR departments: Getting out of the middle. In *HRM in Virtual Organizations*. Edited by Richard L. Henneman and David B. Greenberger. Greenwich: Information Age Publishing, pp. 81–101.

Stolt, Roland. 2010. *Differences and Similarities between Domestic and International HRM*. Norderstedt: Grin Verlag.

Strohmeier, Stefan. 2007. Research in e-HRM: Review and implications. *HRM Review* 17: 19–37. [CrossRef]

Thathsara, A. D. S., and Jayaranjani Sutha. 2021. Investigating the influence of e-hrm practices on organizational performance: The mediating role of organizational agility (with special reference to financial institution). *International Journal of Engineering and Management Research (IJEMR)* 11: 108. [CrossRef]

Truss, Catherine. 2008. Continuity and change: The role of the HR function in the modern public sector. *Public Administration* 86: 1071–88. [CrossRef]

Ulrich, Dave. 1997. *Human Resource Champions: The Next Agenda for Adding Value and Delivering Results*. Boston: Harvard Business School Press.

Voermans, M., and Marc Van Veldhoven. 2007. Attitude towards E-HRM: An Empirical Study at Philips | Emerald Insight. Attitude towards E-HRM: An Empirical Study at Philips | Emerald Insight. Available online: https://www.emerald.com/insight/content/doi/10.1108/00483480710822418/full/html (accessed on 21 March 2020).

Weerasuriya, J. 2006. Effectiveness of E-HRM Systems in Sri Lanka. Master's thesis, University of Moratuwa, Moratuwa, Sri Lanka. Available online: http://dl.lib.mrt.ac.lk/handle/123/407 (accessed on 18 July 2022).

Zhang, Li, and Hong Wang. 2006. Intelligent information processing in HRM: An implementation case in China. *Export Systems* 23: 356–69. [CrossRef]

MDPI AG
Grosspeteranlage 5
4052 Basel
Switzerland
Tel.: +41 61 683 77 34

MDPI Books Editorial Office
E-mail: books@mdpi.com
www.mdpi.com/books

MDPI

Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-2466-3