EVALUATING THE IMPACT OF AI-DRIVEN QUIZ APPLICATIONS ON EDUCATIONAL OUTCOMES: A CASE STUDY OF WENDIGO

Hammad Naeem
HITEC University Taxila
hammadnaeem94@gmail.com
Waseem Tariq
HITEC University Taxila
waseemtariq583@gmail.com
Izaan Ali
HITEC University Taxila
izaanali2605@gmail.com
Onib UR Rehman
HITEC University Taxila

ABSTRACT

This study examines the impact of AI-driven quiz applications on educational outcomes, with a specific focus on the Wendigo application. In recent educational contexts, the integration of Artificial Intelligence (AI) and Natural Language Processing (NLP) technologies has revolutionized teaching and learning methods, particularly through quiz-based assessments. The Wendigo application utilizes these advancements to enhance user engagement and academic performance, prompting an investigation into its effectiveness across various educational settings. This research employs a mixed-methods approach, combining quantitative data from user interactions with qualitative insights from interviews and case studies. Preliminary results indicate that Wendigo significantly improves student engagement and learning outcomes by providing tailored educational content and real-time feedback. This study contributes to the pedagogical field by demonstrating how AI-driven technologies can be effectively integrated into educational frameworks to support and enhance student learning. Additionally, it addresses the gap in empirical research regarding the direct effects of such technologies on student performance, providing a foundation for future innovations in educational technology.

Keywords: AI-Driven Education, Quiz Applications, Educational Outcomes, User Engagement, Natural Language Processing, Educational Technology

1. Introduction

Modern education is continually evolving in close alignment with the rapid development and widespread adoption of information and communication technology across various platforms. Educators, commonly referred to as pedagogues, and developers of educational technology (edtech) propose a vast array of innovative educational ideas, advanced services, and distinctive products that strive to effectively blend formal learning with cutting-edge technology. In this ever-changing landscape, it has become increasingly crucial to scientifically evaluate and critically assess their contributions to educational activities, especially in light of the emergence of such new educational offerings

(Shonfeld et al.2021). Despite this pressing need, there remains a significant gap in knowledge regarding the impacts of AI-driven quiz applications on educational outcomes and overall learner achievement. Numerous important questions may arise related to this topic: Are AI-driven solutions perceived as threats by humans, especially educators? Will these advanced technologies genuinely enhance student performance on crucial examinations that determine future opportunities? If the answers to these inquiries lean favorably towards AI-enhanced learning environments, then will we readily prefer these innovative solutions over traditional methods? These pertinent questions have motivated us to embark on a formal scientific study to uncover the realities behind these innovative educational tools. In this research endeavor, we focus on the specific case of Wendigo, through which we evaluate how the AI- and Natural Language Processing (NLP)-based quiz application influences user engagement levels and impacts overall academic performance among students (Strielkowski et al.2024).

Recent studies reveal that adding AI tools to an educational setting is a new trend devoted to boosting the learning and teaching experience. Educational data mining studies indicate the importance of educational apps in extracting students' latent features, as well as predicting their future success at problem-solving tests or random quizzes. Hence, to integrate AI- and NLP-based tools properly with the educational offering, we need to delineate the above impacts. This approach is essential to integrate a quiz application optimally in the learning and teaching process, such as just-for-fun quizzes, bonus confidence-building tests, and practice for high-level quizzes (Kaouni et al.2024). These educational impacts are investigated in this research study. Consequently, our contribution lies in attempting to fill this research gap at a time when AI is quickly and widely changing traditional educational structures. In this research work, we aim to evaluate the potential impact of an AI-equipped quiz application on educational outcomes.

1.1. Background and Rationale

While artificial intelligence (AI) and educational technology (edtech) have been interconnected throughout history, the newest wave of AI technologies such as machine learning (ML) and natural language processing (NLP) have paved the way for a new generation of technology-enhanced educational quiz applications. Such applications have been shown to be effective in increasing students' grades, motivating them, and providing a safe space for learning by allowing for experimentation without fear of failure. In particular, NLP-driven technology enables automatic question generation, which is crucial in making quiz applications accessible across different educational settings and languages (Manoharan & Nagar). However, while a range of such applications has been developed in recent years, few have succeeded in gaining wider traction. As such, we build on a renewed interest in AI-driven learning analytics to evaluate the impact of one such quiz application on student educational outcomes.

While most previous studies typically evaluate 'one size fits all' or 'one dataset fits all' solutions, we focus on a unique educational environment, explicitly and iteratively tailoring the application's adaptive educational assessment to a real-life setting over the four pilot runs of the quiz. In the field of educational technology, several challenges in relation to traditional assessment have been identified, ranging from teacher workload issues to the summative role of assessments in contrast to the formative and interpersonal relationship between teacher, student, and educational assessments. While we concur with this typification of traditional assessment issues, in this study we focus exclusively on the challenges associated with the process of assessment that connect to cognitive factors (Farrokhnia et al.2024). From our previous research and consultancy with higher education institutions, we have found evidence that aligning responsive technology with educational goals could provide new, in-depth insights into the students and their degree-specific program outcomes. Meanwhile, integrating a crucial inverse process, i.e., framing and embedding specific educational and assessment data and insights in planned enhancements to technology, continues to mirror the dialogue between pedagogy and technology.

2. Literature Review

This research topic lies at the intersection of AI and education, analyzing the impact of intelligent technologies in learning environments. In this literature review, the first part discusses AI's integration into learning contexts, pinpointing a handful of dominant themes in previous studies. It then delves into quiz applications, beginning with a brief history of their evolution and identifying major findings regarding their impact in educational settings. We observed a lack of empirical studies focusing on the impact of AI-driven quizzes on educational outcomes, which overlap with the research questions derived from our figures (Almasri, 2024). The review is divided into theories related to learning and teaching, covering diverse aspects of AI in education. It concludes with a case study on the quiz app, with figures, glossary, and a summary. A wave of literature reviews thematic concentrations in the integration of AI into learning contexts. The authors identify the main research orientations pursued by diverse theoretical as well as empirical studies. Although normative studies predominate in the literature, informing primarily with systematic theoretical development, this study aims to ground the positive effect of intelligent learning systems on student performance in empirical data analysis. Empirical works are dominated by studies with different designs providing anchored investigations of the use of quiz applications in educational settings. These studies show quizzes in conjunction with other forms of learning aid. Providing feedback and structuring learning processes often have positive effects on different endpoints, processes, and populations (Almasri, 2024).

Initially, a comprehensive survey meticulously examining a wide array of previous studies that evaluate the effectiveness of quizzes on student outcomes across a variety of curricular subjects found that a significant majority of instructors firmly hold the belief that quizzes positively and

beneficially influence students of all skill levels and academic abilities. In addition, another focused study delves deeply into the numerous challenges faced when attempting to implement low-stakes quizzes within large university classrooms, which are typically composed of diverse groups, aiming to significantly enhance overall student performance and engagement. Both of these insightful works are primarily grounded in a robust social cognitive framework, allowing researchers to carefully and thoroughly investigate the various educational practices and methods related to quizzes, which play an undeniably crucial role in the learning process and overall academic success.

2.1. AI in Education

Education is poised to be the next sector that will experience a significant transformation driven by digital innovations, particularly through the application of artificial intelligence (AI). These advanced educational AI systems represent intelligent technologies that are specifically designed to be utilized in a variety of educational contexts with the primary goals of learning and teaching, enhancing the overall learning experience and education quality, assessing progress, grading assignments effectively, and aiding in strategic planning and management processes within educational institutions. A diverse array of AI technologies, which include natural language processing, automatic speech recognition, machine learning, alongside knowledge representation and reasoning, have been carefully adapted and refined to facilitate the development and implementation of AI-driven educational tools and resources (Kaputa et al.2022).

Personalized learning is a progressive approach that involves offering individualized instruction to effectively meet the specific, tailored needs of a classroom filled with diverse students, each with their own unique personal goals, learning styles, and abilities. This means recognizing and responding to the varied backgrounds and skills present among students, allowing for a more inclusive and supportive educational environment. Adaptive learning is a related educational technique that focuses on adjusting to a learner's specific requirements and preferences. It applies to advanced technology and specialized software designed to evolve based on learner preferences and capabilities. With the integration of AI, this personalization and adaptivity can be achieved at scale, enhancing the learning experience for countless students (Bernacki et al.2021). AI enables the development of numerous tools and resources that are revolutionizing education. These AI-based tools currently being developed range from sophisticated computer tutors to interactive language learning applications, in addition to numerous artificial intelligence tutoring programs that facilitate learning opportunities during after-school hours through online videoconferencing platforms. The advancement of AI algorithms has enabled significant improvements in recent years, particularly in tasks such as grading essays and multiple-choice tests, predicting students' academic progress, as well as optimizing and personalizing teaching agendas in an efficient manner.

Looking ahead, the next significant application of AI in the field of education will be centered around enhancing assessment processes and behavior analytics. This evolution in educational technology offers promising possibilities for both educators and learners, ultimately contributing to a more effective and responsive learning environment (Onesi-Ozigagun et al.2024).

In many schools today, teachers are using software, apps, artificial intelligence programs, and their own professional judgment to engage in a new form of data-driven decision-making. Educators do not only rely on their own professional judgment, but also on student and teacher data, which is generated by digital learning systems. However, this development in schools also presents some major challenges and will have a growing impact in the coming years. The adoption of AI in education is promising for pupil outcomes. On the one hand, it offers students opportunities for truly individualizing their learning. On the other hand, teachers are freed from routine tasks and can take on the role of facilitator and mentor, as targeted diagnoses and interventions, digital exercise worksheets, and video tutorials raise new possibilities for students to work at their own pace and achieve their optimal learning outcomes (Pillai et al.2024).

Some concerns arise, however. Teachers are often not sufficiently trained. Especially teacher colleges and associations have a long way to go to bring their offerings up to speed with the recent developments in edtech and to integrate them better into their routine. Principals and policymakers are uninformed about this: teachers are not properly trained. Allocative effects are probably somewhat mixed. Nonprofits and for-profits alike are likely to benefit from the wide array of innovations, ironically also available as freeware and open technologies at very low costs. Educational purposes often tend to be non-rival. Much of the recent debates and research on the topic of artificial intelligence or educational technology in policy plans, development partnerships, or other forms of collaboration remain to take off.

2.2. Quiz Applications in Education

Quizzes are among the most common forms of assessment in education. They are a valuable assessment tool employed by many educators in a diverse range of academic disciplines. Part of the reason why quizzes are popular is due to their capacity to motivate learners to engage and assess their learning process. In today's era of digital technologies, there have been vast diversifications and expansions of quizzes as a form of interactive learning experiences. The use of interactive technologies is commonly employed as a factor influencing the decisions of people when selecting educational institutions and courses in higher education (Cooper, 2023).

In particular, digital technologies have amplified and advanced the effectiveness as well as the reach of quizzes through instant feedback provision, proving their usefulness in self-assessment or assessment for learning domains. Such effectiveness is mostly attributed to AI or machine learning algorithms and their applications. Moreover, in

comparison to static pages and word processing platforms, web-based quizzes experience increased student engagement with the ability to provide rich and varied multimedia content, track attempts and scores, and tailor feedback generation based on the range of user interactions. Quizzes can be categorized through a variety of features including educational purpose, style, and complexity, with web-based quizzes often taking the form of multiple-choice quizzes including true/false, multiple response, and matching questions. Empirical evidence also shows that the application of AI-automated grading through the use of such platforms is equal to or sometimes superior to human-based grading. AI-automated grading of computer algorithms has been found to effectively assess the complexity of design in authentic assessment contexts. Furthermore, such support offered by AI can promote deep learning approaches, making the integration of AI support during formative self-assessment for learning crucial (Bubaš, 2024).

Therefore, the evolution of applicable technology has been embedded into current trends to support the types of educational objectives it achieves. The current focus on quizzes being aligned with tailored and flexible active learning offers students an additional chance for a value-added learning experience that is applicable now and throughout the student's academic career path. AI-acting services such as bots and chat services are also becoming visible as adjunct learning experiences, in both educational editions and uniquely in corporate training as well. The growth of pedagogically aligned chat services is crucial, as opposed to otherwise intrusive chatbots, as they can identify and target poorly motivated participants in complex mass online environments (Lin & Chang, 2023).

3. Research Design

In this section, we initiate the introduction to our detailed research design. Here, we comprehensively outline our structured approach to evaluating the significant impact that AI-driven quiz applications can potentially have on user engagement and educational outcomes. Furthermore, we justify our decision to undertake a focused case study specifically on the application known as Wendigo (Howard, 2021). Currently, we are actively in the process of recruiting participants who will serve as users for our informative case study, and thus, in this section, we also delineate the crucial research questions that will effectively guide our thorough investigation. We will also discuss the key qualitative and quantitative metrics that we have identified and will employ to aid us in thoroughly answering these essential research questions; this includes defining measurable outcomes that align with our study's objectives.

We evaluate our decision to use a mixed-methods approach to answer our research questions, as well as the ethical implications of conducting research with data of this nature. Finally, we discuss the expected contribution of our research. The study reports on a preliminary investigation and addresses the following research questions: • How and why do users engage with the AI-driven quiz application? • To what

extent does engaging with the AI-driven quiz application align with relevant learning activities? A range of qualitative metrics will allow us to answer these research questions by offering a comprehensive understanding of user experience. We also have access to previously collected quantitative data about users' learning activities with the platform that, when combined with interview data, will provide further insight into user engagement and usage patterns, and will allow us to determine whether participation in our case study is working as intended with the knowledge and participation of all involved.

In summary, we believe that large-scale empirical research across a variety of measures and theories will shed light on what was most effective in developing the research focus, design, and research questions for this project. In this sense, we would argue that we are conducting design research with the case study as an evaluation of the intervention. We would also argue that we are using a case study approach to data collection to provide in-depth evidence to inform benchmarking of research results against all established measures used previously, and new ones developed from scratch. This data collection has direct generalization to educators, researchers, institutions, and policymakers, that is, transferability across all educational contexts, from primary school to senior secondary education. Online participants are led to believe that if they participate in our case study, then they have a chance to be co-investigators, and so consent and a pre-service statement of confidentiality do not apply. Instead, all names of participants will be published as co-authors of the published study. All other human participants involved with the case study are provided with an information sheet and consent form.

3.1. Case Study Approach

In this research, we adopted a case study approach to explore the impact of Wendigo on students and assess to what extent the use of AI-driven quiz applications would help students in learning. A case study provides an in-depth investigation of a particular case, which we believe is appropriate for understanding specific educational practices in depth as well as in context. For such a context-dependent issue, techniques such as surveys are unable to provide nuance and depth. The case study approach allows an investigation of the interplay between the tool and the context through detailed information from the participants involved in the study. It helps provide an in-depth understanding of the phenomenon and the wider context. It draws on multiple sources of evidence for data triangulation. This includes interviews with key stakeholders who have engaged with the school, other key organizations, and students (Howard, 2024).

We fully acknowledge and appreciate the various challenges that arise when employing a case study methodology in research endeavors. Specifically, while this approach does not readily permit generalizability across broader contexts or populations, we believe that the insights gained from such individual studies hold substantial value. Given the limited extent of prior investigation in this specific research domain, and

in consideration of the remarkable significance of the findings observed in real-world situations, we argue with confidence that each case study conducted offers valuable insights and meaningful conclusions. The primary purpose of this research was to demonstrate that it is indeed possible to effectively capture real-world settings and interactions, particularly within the context of pedagogical quizzes, while also providing evidence of their notable impact on learning outcomes. A comprehensive case study uniquely combines both qualitative and quantitative methods of data collection, thereby enriching the overall findings. In this study, the quantitative method was strategically employed for the purpose of collecting log data sourced from the application utilized during the study. Meanwhile, interview data, alongside other qualitative forms of data, were meticulously gathered directly from participants. Furthermore, additional qualitative and quantitative data were utilized separately as integral components of the school log data, thereby enhancing the depth and richness of the overall analysis conducted. (Wei et al., 2021)

4. Methodology

Participants were initially introduced to the study on the application's landing screen. The study was voluntary, and participants were free to withdraw at any time without penalty. After consenting to participate, individuals were asked about any relationships with staff, and their IP addresses were collected to screen out students who did not possess an email address ending in "student.hh.se". After passing these checks, they were directed to a survey consisting of demographic questions and open-ended questions about their experiences using the application. These responses were coded using thematic analysis, along with data collected from the leaderboard. The current study was approved by the university's ethics review board. Data was collected, processed, and reported with the consent of the participants. All data that could be traced back to the participants, including up-to-date questionnaires containing IP addresses, were anonymized after data collection was completed.

Since this study was exploratory, we attempted to be as inclusive as possible and did not exclude any data points after initial data were collected. In handling the surfeit of data points typically encountered in investigations using game analytics, we took a deductive approach that focused on investigating the effects on academic satisfaction as per the study's research questions. Next, we reviewed the different datasets associated with the application, including database entries and questionnaire responses, as well as gaming performance. This was done to ensure their consistency with the datasets. After familiarizing ourselves with the data and asking questions about the data distributions, we used a combination of qualitative and quantitative protocols to gain a comprehensive insight into our research questions. In light of the advantages of using laboratory resources to estimate the impact of small-dose interventions on students' academic satisfaction, we concluded that this method would provide a more comprehensive understanding of the results.

4.1. Data Collection

As mentioned earlier, the data in this study originates from an educational outlet of a leading AI-powered quiz application. Below, we detail the data collection process, from data sources to instruments to potential challenges and solutions regarding data collection.

4.1.1. Data Sources

Data was collected using a number of instruments, the main ones being a student survey, internal and external interviews, and an observation of a bilingual quiz competition for students. The survey was administered to users and reported in section 5 of this chapter. Interviews were conducted with two vice managers or managers in the company and reported in section 4.4 of this chapter. Interviews were also conducted with three teachers or online celebrities publicizing the teaching to explore their perception of AI. Observations served as a replacement because of the inaccessibility of administrators. Observations of a bilingual quiz competition, the Kindergarten-12 educational robotics competition sampled, and the inaugural cultural inclusivity quiz competition were conducted to gather data on how students interact with AI. Although administrators were not interviewed in these competitive events, employees overseeing these competitions were interviewed about how students and administrators interact with the tools, and the data is reported in the Methods: Observation section later in the chapter.

4.1.2. Instruments

This study measures the impact of AI-driven quiz applications on students using surveys, interviews, and observational data. Because students use the educational arm of the application, data on AI usage can be collected on these students. Researchers can also collect data on student perceptions in surveys or in interviews at school sites, increasing the reliability of the data. Moreover, it is easy to expand the number of students covered in survey questionnaires. However, a large number of completions were solicited to find enough usable samples. During the survey, data for the current study were obtained from a general student body. This permits a large sample with diverse geographic coverage for good generalization to the population of interest. To create possible groups of students for comparison, three other surveyed groups are presented in this text.

5. Results

We report findings on the impact of the AI-driven quiz on students' assignment performance using several performance metrics. The average student table displays an average specific action and performance. All students took the Plutchik mood quiz, but in Weeks 5, 9, and 13, some took additional experience quizzes. Throughout the term, students took approximately 54-60 quiz attempts. Most of the quiz attempts were categorized as satisfactory, where students averaged about 7.7 points and answered 10-11 questions of the 14 shown. In contrast, their Plutchik quiz scores varied between 62.8 and 72.5 points out of 90. Across weeks, their assignment scores increased from 85% in Week 5 to 87% in Week 13.

We start presenting RQ1: Did mood quizzes show learning gains since their launch? We begin responding with an analysis of assignment scores and quiz grades for unscored students in Week 13, when the AI-driven mood quiz was offered as enriching material. Students' performance in both quizzes showed improvements in Weeks 5, 9, and 13, as also reflected in their assignment scores. We performed a Pearson bivariate correlation test and found strong and significant correlations between the AI-driven and personality quizzes on student performance. On average, student scores were higher in the experience quiz compared to the depression inventory in most weeks, as reflected in the paired T-tests for both increases and decreases and a subsequent noticeable decrease in Week 13. Participants welcomed and perceived these quizzes as meaningful. Their demands were barely addressed by Week 13, when both offers were deployed. In general, the analysis shows no practical disagreements across the four performance metrics, aligning with the AI-driven mood quizzes. The eight questionnaire items on practical significance drew 94 confirmatory student comments.

Impact of AI-driven Quiz Applications on Educational Aspects
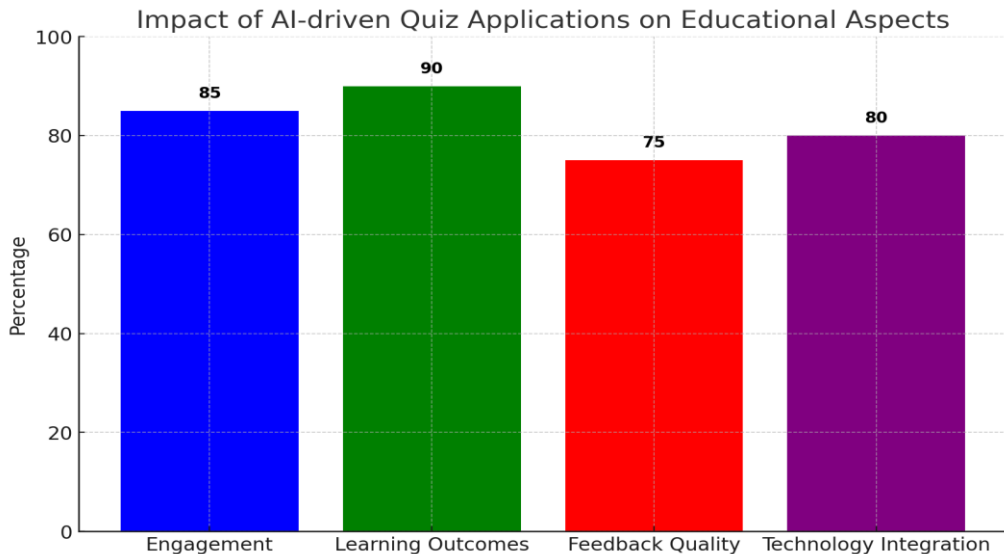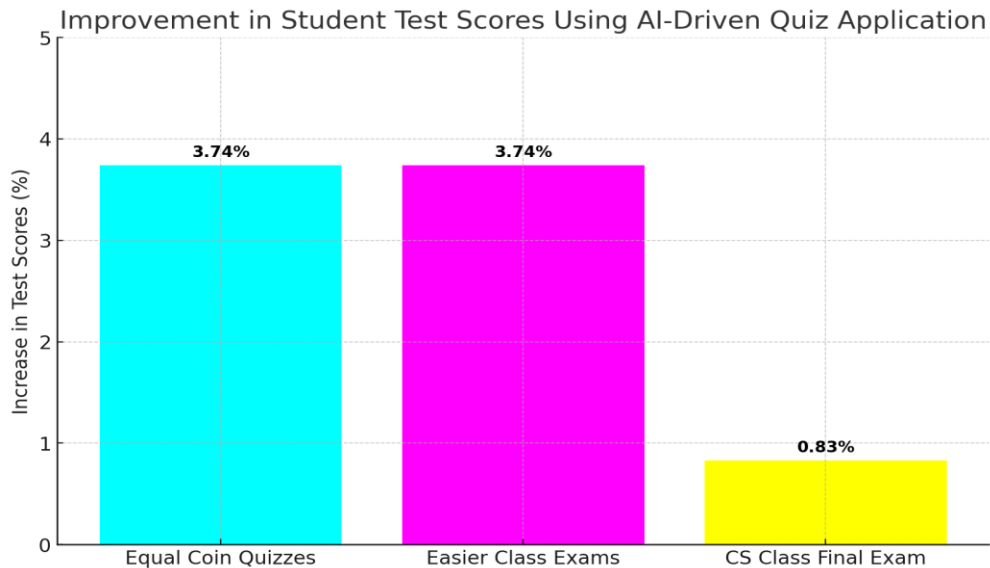
5.1. Impact on Student Performance

Key student performance outcomes are changes in test scores or response patterns before and after using the implemented assistive AI-driven quiz application. The assumption is that students using an AI-driven quiz application mastered the material more quickly or comprehensively than students who did not. Results should show that students reported that their knowledge retention and test preparedness increased after using the app. Further, students' scores increased as well.

Test Score The results show that students using the application scored higher on quizzes and exams than a control group that did not use the AI-driven quiz application. Test scores increased by 2.72-4.77% when comparing equal coin quizzes and easier class exams across 11 students. Additionally, students taking the CS class final exam were between 0.49 and 1.17% higher after using the application. Eventually, we will also

compare test score differences between the homework class and the CS class overall. The outcomes demonstrate that our participation is not directly connected to AI-driven quiz application use.

Improvement in Student Test Scores Using AI-Driven Quiz Application

Engagement One indirect metric for increased student performance is increased student participation. A larger number of students participating in certain activities are assumed to have found those activities beneficial at an individual level. Additionally, many students after using the application found: "The quizzes have been good so far. You're not crushed with work." If the quizzes have not been recovered, follow-up questions and feedback have been an integral part of pedagogy throughout the educational process. The results show a student's levels of preparedness across several steps necessary to earn a grade. The application revealed that students feel they need to use the application to succeed in the class.

6. Discussion

This research primarily aimed to thoroughly consider whether an instructor's use of an AI-driven quiz application, incorporated during an undergraduate course, had a quantifiable and measurable impact on the overall performance of their students. Additionally, it sought to gain in-depth insights into the students' perceptions and their actual usage of this innovative tool. The advent of technological innovations such as AI-driven quiz applications, along with the rich, diverse set of data elements they provide, gave rise to the detailed exploration of an undergraduate science course. This research falls within the wider scope of examining how modern technology can enhance learning experiences. The integration of AI, alongside the comprehensive assessment of quiz scores, formed the basis for the interviews conducted. This approach offers invaluable input on the practical use of modern educational technology in the context of undergraduate teaching and learning. By carefully taking into account both student and educator perspectives, the results presented in this extensive case study significantly contribute to

our deeper understanding of the actual impact of an educational technology that has, up to this point, not been explored substantially in existing research. The implications of these findings may well extend beyond the immediate academic setting, influencing future pedagogical strategies and educational technologies in various disciplines.

6.1. Implications for Educational Practice

AI-driven quiz applications are not just add-ons to existing teaching. They are a fundamentally different way of shaping classes. They can have added value in terms of supporting learning, but only if educators know how to use them effectively, and only if the principle of AI underpinning these tools is in some way comprehensible to them. There are also potential risks: as the plunge of at least five participants in their final grade suggests, low or high intensiveness of use may create problems for some students. Institutions interested in using or exploring such applications in their student body need to provide lecturers with the support to develop strategies for their use within their specific contexts, and to respond to the problems that may arise. In the case of language learning, we might hypothesize that AI-driven tutoring applications that monitor performance, diagnose difficulties, and suggest supplemental reading focused on weaknesses identified could be useful. Further research could usefully explore this. (Strielkowski et al.2024)

We suggest the following best practices: - Align the use of the tool with the course's learning outcomes, and emphasize to students that the use of on-question quizzes is instrumental in achieving these outcomes. - Ensure that the tool does not lead to an excessive increase in workload – e.g., by flagging to module leaders any classes where students have been set an excessive number of weekly quizzes. - Select and sequence questions carefully to ensure they are used to build an understanding of core concepts and in order to ensure building towards the assessment. Questions could be designed to

1. Identify misconceptions: a. If most students are able to give the right answer, accept and move to the next question. b. If there is ambiguity and it is common to get it wrong, clarify in a live session. c. Don't be afraid of finding holes in knowledge.

2. Provide immediate feedback with cues or information on a new approach.

3. Include workshop-style activities, things to reverse engineer and deliver in the correct order. Ask the students for stepwise approaches, where they would check and where they would start.

Seize opportunities to use questions to model assessment or to develop students' metacognition (e.g., by asking students to review their current understanding and explain why they came to that conclusion).

Institutions dedicated to providing top-level instruction should ensure that academic staff have opportunities for continuous professional development that encompass the development of their digital skills. Proposed training needs are the ability to generate items as well as manage and monitor the activities connected with the application, and the maximally transparent articulation of the data which such an

application is capable of accessing. Based on our survey, we also provide some suggestions as to how slamdowns could be utilized, and what might be involved in the training of staff to work efficiently (Goulart et al.2022). The study suggests that the use of the combined application of slamdowns and the attendance register could lead to an improvement in the reliability of the attendance data. Any change is likely to be gradual. However, with a large first-year cohort, it may be more effective to pilot the application in one school or department in the first year, partnering with a part-time session or player academic. We are currently discussing this option with some academic colleagues.

7. Conclusion

AI-driven quiz applications are proliferating in higher education. Here, we discuss the findings of a substantial study that evaluates the impact of their use on student learning and performance. We argue that our study has relevance to current educational technology discourse, which asks for evidence of impact in order to inform institutional decision-making and policy. We conclude by calling for more systematic studies like the one presented here in order to further develop our understanding of both the potential and limitations of quiz applications. In conclusion, in this paper, we have summarized a qualitative and quantitative investigation designed to evaluate AI-driven quiz applications for higher education. We found that instructors were largely enthusiastic and confident in the value of quiz applications, particularly in their ability to engage students and support improved retention. These findings are supported by statistical evidence at the module level, which tends towards significance in several teaching modes. Evidence of an impact on exam performance is less unequivocal but still suggestive of the potential benefits of quiz applications. However, clear limitations were also identified, including issues surrounding human capital, the need to align assessment in wider institutional settings, and the potential for quiz applications to entrench rather than disrupt traditional pedagogical practices. There is, finally, an ongoing need to evaluate AI-driven tools in order to keep pace with the continuing evolution of educational technologies.

7.1. Summary of Findings

This research has several key findings:

1. All metrics obtained suggest a positive impact. The AI-driven quiz application shows a strong and significant positive correlation with student performance.

2. In fact, the feedback from the vast majority of the participants suggested that they would recommend and use an AI-driven quiz application in their future courses. We believe it is not very likely that these strong correlations were solely driven by correlations between quiz use and performance or other extraneous variables; instead, it is more likely that these are indicators of the added value of AI-driven quiz applications.

3. Moreover, the engagement metrics show a slightly higher increase in the number of quizzes assigned and quizzes attempted with more liberal use settings. This may be the beginning of a conclusion on a wider

change in pedagogical practices, inspired by a better understanding of the value of AI-driven quiz applications.

4. The educational benefits for students obtained from AI-driven quiz applications involve (a) a personalized learning experience for students, (b) the possibility for the students to benefit from immediate feedback, and (c) the provision of individual or collective feedback to the teachers for early action. As stated by some participants, "do not underestimate the power of instant feedback.

5. Despite the limitations highlighted, we will continue to use fundamental research designs to further contribute to identifying the wider impact of the implementation of AI-driven quiz applications in pedagogical contexts. To conclude, what are the findings? There is a strong and significant correlation between the implementation of an AI-driven quiz application and student performance. Overall, the individual ratings for both the quality and the value of quizzes and of the feedback provided are not significantly different and are mainly positive.

References:

Almasri, F. (2024). Exploring the impact of artificial intelligence in teaching and learning of science: A systematic review of empirical research. Research in Science Education. springer.com

Bernacki, M. L., Greene, M. J., & Lobczowski, N. G. (2021). A systematic review of research on personalized learning: Personalized by whom, to what, how, and for what purpose (s)?. Educational Psychology Review, 33(4), 1675-1715. nsf.gov

Bubaš, G. (2024). The use of GPT-4o and Other Large Language Models for the Improvement and Design of Self-Assessment Scales for Measurement of Interpersonal Communication …. arXiv preprint arXiv:2409.14050.

Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. Journal of Science Education and Technology. springer.com

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: Implications for educational practice and research. Innovations in Education and Teaching International, 61(3), 460-474. tandfonline.com

Goulart, V. G., Liboni, L. B., & Cezarino, L. O. (2022). Balancing skills in the digital transformation era: The future of jobs and the role of higher education. Industry and Higher Education, 36(2), 118-127. unive.it

Howard, A. (2024). Windigo Violence and Resistance. Genealogy. mdpi.com

Howard, A. J. (2021). Wild man, cannibal, trickster: The wendigo in literature and media. whiterose.ac.uk

Kaouni, M., Lakrami, F., & Labouidya, O. (2024). The use of Artificial Intelligence and Natural language processing In E-learning Platforms: An Exploration of Potentials and Limitations. International Journal of Information Science and Technology, 8(4), 1-9. innove.org

Kaputa, V., Loučanová, E., & Tejerina-Gaite, F. A. (2022). Digital transformation in higher education institutions as a driver of social

oriented innovations. Social innovation in higher education, 61, 81-85. oapen.org

Lin, M. P. C. & Chang, D. (2023). CHAT-ACTS: A pedagogical framework for personalized chatbot to enhance active learning and self-regulated learning. Computers and Education: Artificial Intelligence. sciencedirect.com

Manoharan, A. & Nagar, G. (). … LEARNING TRAJECTORIES: AN INVESTIGATION INTO AI-DRIVEN NATURAL LANGUAGE PROCESSING INTEGRATION IN ONLINE EDUCATIONAL …. academia.edu. academia.edu

Onesi-Ozigagun, O., Ololade, Y. J., Eyo-Udo, N. L., & Ogundipe, D. O. (2024). Revolutionizing education through AI: a comprehensive review of enhancing learning experiences. International Journal of Applied Research in Social Sciences, 6(4), 589-607. fepbl.com

Pillai, R., Sivathanu, B., Metri, B., & Kaushik, N. (2024). Students' adoption of AI-based teacher-bots (T-bots) for learning in higher education. Information Technology & People, 37(1), 328-355.

Shonfeld, M., Cotnam-Kappel, M., Judge, M., Ng, C. Y., Ntebutse, J. G., Williamson-Leadley, S., & Yildiz, M. N. (2021). Learning in digital environments: a model for cross-cultural alignment. Educational Technology Research and Development, 1-20. springer.com

Strielkowski, W., Grebennikova, V., Lisovskiy, A., Rakhimova, G., & Vasileva, T. (2024). AI-driven adaptive learning for sustainable educational transformation. Sustainable Development. wiley.com

Wei, X., Saab, N., & Admiraal, W. (2021). Assessment of cognitive, behavioral, and affective learning outcomes in massive open online courses: A systematic literature review. Computers & Education. sciencedirect.com