ONE-TO-ONE CONVERSATIONAL STUDY PARTNERS FOR CHILDREN USING AI: A USER EXPERIENCE ANALYSIS OF TEDDYAI

P. Otermans, A. Sharma, M. Singh, D. Aditya

Otermans Institute, Brunel University London (UNITED KINGDOM)

Abstract

AI-driven educational tools can be effective for children's learning experience, and especially conversational AI can have a positive impact on a child's learning journey. Using conversational AI, with generative AI language models, personalised study partners can be made for children that can learn with them, motivate them, teach them, and gamify their learning experience using one-to-one conversational interactions. In this study, TeddyAI, a conversational AI learning app, tests this concept. It is a conversational Teddy bear, powered completely by AI, that supports learners between the ages of 4 and 7. From a teaching perspective, it helps them learn alphabets, and topics like general geography, history, animals in a zoo. Apart from these, TeddyAI is also able to teach two subjects from Key Stage 1 and 2 of the UK national curriculum. The objective of this research is to understand the user experience and overall usability of TeddyAI to offer a personalised approach to learning by delivering tailored experiences through the study partner; Teddy. To understand the user experience of TeddyAI, 35 primary school children interacted with TeddyAI, each for 12-15 minutes. During their interaction, their real-time engagement was measured, and they were observed. The System Usability Scale (SUS) was used to gather user feedback and was followed by a discussion to explore the experiences of the children in detail. Results showed that the SUS score was 72.6 indicating a score above the average of 68. This suggests that TeddyAI's usability is good, the children found it easy to play TeddyAI, and loved their interactions with him. This study provides insights for future improvements for AI study partners, focusing primarily on user interaction and engagement.

Keywords: Artificial Intelligence, Conversational AI, Education, Personalised learning, Generative AI.

1 INTRODUCTION

The concept of instructional strategy has always been a part of teaching involving designing and implementing a well-structured lesson plan that caters to the diverse learning needs and styles of students, ultimately optimizing their understanding and retention of the subject matter. To achieve the set objectives teachers utilise a variety of teaching methods, such as direct instruction, questioning, guided practice, and collaborative learning, to engage students actively and foster a deep understanding of the concepts being taught. Throughout the session, it is essential to provide opportunities for formative assessment, allowing students to demonstrate their understanding and receive immediate feedback from the teacher, while also informing the educator about areas that may need further clarification or emphasis. A well-designed instructional strategy emphasizes critical thinking and problem-solving skills [1]. In addition, encouraging students to apply their knowledge to real-world contexts and explore connections between different subject areas help foster a comprehensive understanding and support a lifelong love of learning [2]. Consequently, it becomes necessary for a teacher to plan an individual specific learning strategy which seems rather impractical. As technology is advancing expeditiously, there has been a growing need for educational tools that effectively engage and educate children, seamlessly blending into their daily routines. Given the prevalence of smartphones and tablets in our lives, developers and researchers have been focusing on creating intuitive applications, such as educational games and virtual assistants, that address this need. The widespread acquisition of these devices offers an opportunity to explore new ways of learning and catering to the diverse learning styles of children. Furthermore, the integration of AI technologies in digital learning tools holds promising potential for adaptive, personalized learning experiences [3, 4]. The use of AI and machine learning algorithms enables educational tools adapting to individual learners' needs and preferences, better supporting their learning process [5]. For instance, they can analyse a student's performance on multiple assignments and these areas can be reckoned where the learner struggles or excels. Then, AI can tailor future assignments to focus on areas of weakness, helping the learner to improve in these specific domains, or provide more challenging tasks for areas they excel in to further their knowledge.

Teddy AI is one such innovative application designed to enhance the learning experiences for children at an average age of 8 years old. By integrating Artificial Intelligence (AI) with an interactive, game-based format, Teddy AI aims to create a captivating environment for children to learn and practice essential concepts such as alphabets, recognition of objects, and subject-based quizzes. This innovative combination of technologies is of particular interest, as there is growing evidence supporting the role of technology-enhanced learning in increasing engagement and learning outcomes for children [6]. The current study seeks to examine the user experience of the Teddy AI app among primary school children. Specifically, it aims to assess the ease of use, engagement, and satisfaction experienced by children while interacting with Teddy AI. The study employs the System Usability Scale (SUS) [7] and additional qualitative feedback to analyse the strengths and weaknesses of the app, providing insights into potential improvements that could elevate the learning experience for young children.

Moreover, the interactive nature of the app, coupled with the reward system, played a significant role in creating a positive experience for the children. However, the study also highlighted some areas where refinements could be made to enhance the learning experience further. User experience is a critical aspect to consider in the development and evaluation of any educational tool, as it directly impacts the effectiveness of the learning process [8]. Prior studies have emphasized the significance of various factors, including navigational simplicity, age-appropriate content, and personalized features, in the design of compelling and efficacious learning tools [9]. A tool that is not user-friendly or engaging enough may fail to capture the interest and motivation of its target audience, leading to suboptimal learning outcomes. Hence, this research aims to fully harness the potential of Teddy AI in order to reach the desired objective. Teddy AI app contributes valuable insights to the ongoing development of AI-enabled educational tools catering to the unique needs and expectations of young children. The findings could help inform future updates and refinements to the TeddyAI app and guide educators and developers in creating engaging, pedagogically effective tools that align with the evolving landscape of digital learning. alignment]

2 METHODOLOGY

2.1 Participants

The study involved 36 participants with an average age of 8 years and a standard deviation of 1.3 years who were pupils at a primary school. Due to a technical error, only data of 35 participants was used in the analyses. The participants were recruited by the team of Otermans Institute via contacting schools based on their willingness to participate in the research survey.

2.2 Materials

Each participant individually had approximately 10-15 minutes to play and explore the game (APK version 9) and their interactions with TeddyAI were screen recorded simultaneously (where possible). During this time, the researcher was noting any difficulties, negatives or positives as per the user's experience. After the span of 10-15 minutes, the participants filled out a short survey about their experience and impressions of Teddy AI which employed the SUS [7]. The SUS scale is a Likert scale aimed at assessing 10 statements about usability and reliability where kids could provide feedback by circling the smiley based on their experience. It ranges from 0-4 where 0 stands for strong disagreement and 4 for strong agreement. Three additional Likert-type questions were added to further explore the user's satisfaction. Furthermore, open-ended follow up questions were asked to gain better understanding of the user experience about Teddy AI.

2.3 Procedure

Participants were asked to take part in a game-based study which lasted for approximately 10-15 minutes and was easily accessible via smartphone or tablet. The participants were asked turn by turn to play games and explore the Teddy AI under the observation of a researcher from the team of Otermans Institute. Initially, all the participants were asked to open the Teddy AI app and the whole user experience was screen recorded (where possible). The app asked for the audio permission before progressing towards the game to capture the user interaction with the software. An AI-modelled like a Teddy appeared on the screen introducing it as 'Teddy' and asked the user to provide basic details like their name, date of birth, gender, ease of English usage which was easily filled using a keyboard typing method. Then the user had a choice to explore the app as per their convenience by choosing anything between lesson, map and quiz. The lesson tab included delivery of alphabets to participants in which Teddy first showed the letter (A, B, C,) and then said an alphabet along with the word like A is for Apple and asked the user to repeat. If the

participant recognised the word and pronounced it correctly, they were rewarded with coins and verbal appraisal like bravo, well done, superb, great job and upon incorrect repetition they were motivated to try again. Also, then to test the participant's understanding they were shown random pictures based on different themes like food items, games, zoo, etc. and were asked to recognise them. As like previously the same rewarding system was followed up. This feature employed voice recording via google to code the user response each time during the activity. The map tab included different places like the playground, zoo, kitchen, hospital and school which contained learning of the words associated with that place using depiction and repetition method and the user was free to choose any place of their choice for exploration. The quiz tab included various topics like environment, seasons, landmarks, hospitals, geometrics, oceans, relationships and outer space which were made suitable for Key Stage 1 and 2 students. All the topics contained a couple of questions and participants were given four choices per question to choose within 15 seconds. Upon right choice they were rewarded with a few coins and verbal appraisal like bravo, well done, superb and upon wrong choice they were motivated to try again. Finally, a concluding message appeared on the screen depicting their total score and was followed by a feedback box containing a smiley emoticon to record their experience with Teddy AI. Also, the app contained a store where the participant could buy anything from the store like glasses, hat, clothes for Teddy using their earned coins making the overall user experience a lot more engaging. After the exploration of the app, the participants filled out a short survey about their experience and impressions of Teddy AI which employed the SUS. They answered 13 statements on the scale of 0 to 4 based on their level of experience. Some statements were reverse coded due to their negative implications and a final average SUS score for each participant was calculated and recorded.

2.4 Ethical considerations

The study adhered to ethical guidelines for research involving human participants and was approved by Otermans Institute (001-Oct2022). Informed consent was obtained from all participants (either from their parents or via the school), and they were free to withdraw from the study at any time. Confidentiality and anonymity of participants was maintained throughout the study.

3 RESULTS

The participants in this user testability study found that the Teddy AI was above benchmark, set at 68, for the usability of a product. The mean SUS score was 72.6, ranging from 52.5 to 90.9 and the majority of responses feel in the 71.1-81.3 bracket, with 60% of the participants scoring higher than 68 (Fig. 1).

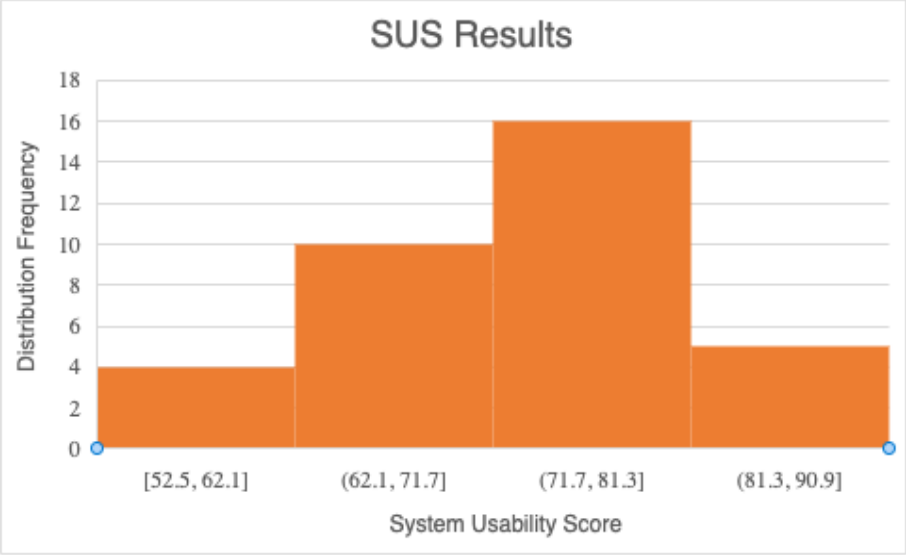

Figure 1. SUS results.

In more detail, the average SUS score awarded for each group is depicted in Fig 2, also indicating a score higher than the benchmark for each age group.

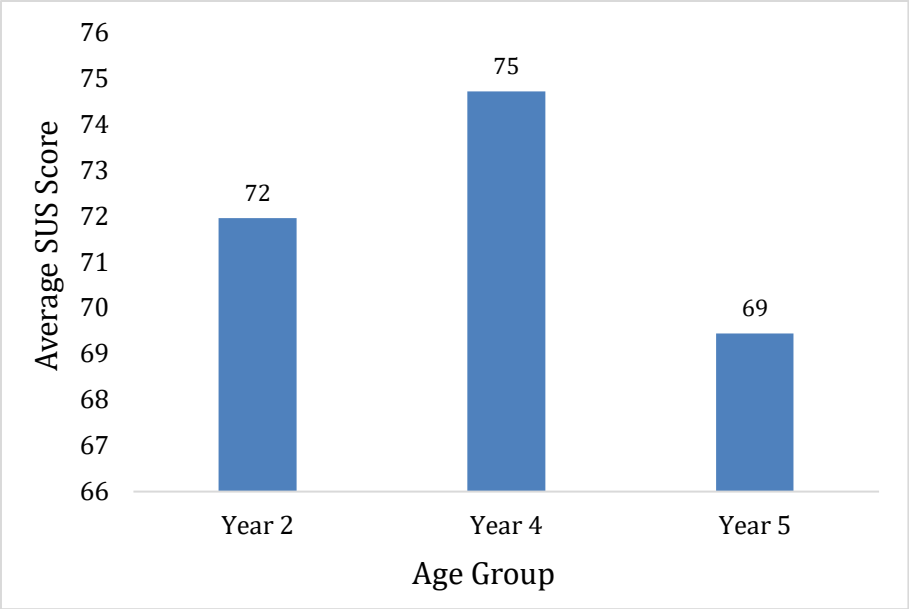

Figure 2. SUS results per year group.

Analyses of the additional questions showed that the children thought playing with Teddy was very fun, many planned on telling their friends about Teddy AI, and if there was more time almost all children wanted to continue playing with Teddy (Fig 3).

The researchers also observed the children when they were interacting with Teddy AI. Children liked the points system and children were keen to improve their score. They also enjoyed the AI feature of Teddy AI, i.e. the freedom of having a conversation. Most children we able to complete all tasks and engage in the conversation with Teddy AI on the app. Children really engaged with the shop feature and wanted all sort of other items. They were very excited about the interactivity and they liked that you can try to go to many places and try different quizzes and activities. They also liked that Teddy gave them feedback. There were also areas that children struggled with. Some needed help with putting in their birthday. For some children navigating the map was not obvious at first, but once they found the swiping features, everything was fine. Some children found the quiz questions too difficult.

Based on the findings of this study, suggestions for future iterations include the following:

- Tailor the difficulty level of the activities to age groups, English ability and other indicators of performance. In other words, easier questions in the quiz for younger age groups, more difficult questions in the locations for older age groups.
- Test knowledge recall; track which questions the learners have already been presented with.
- Young learners may understand inputting age rather than their full date of birth.
- Variety of storylines to assist the gameplay would draw more learners in and do more active things with Teddy.
- Variation of games; puzzles, reaction games, mobility games (actually moving Teddy, connecting dots etc).
- Ability to adjust the pace of the game or Teddy's speech in the settings.
- Add a short intro where Teddy points out the different available features of the app so learners are aware of what features they can access and explore.

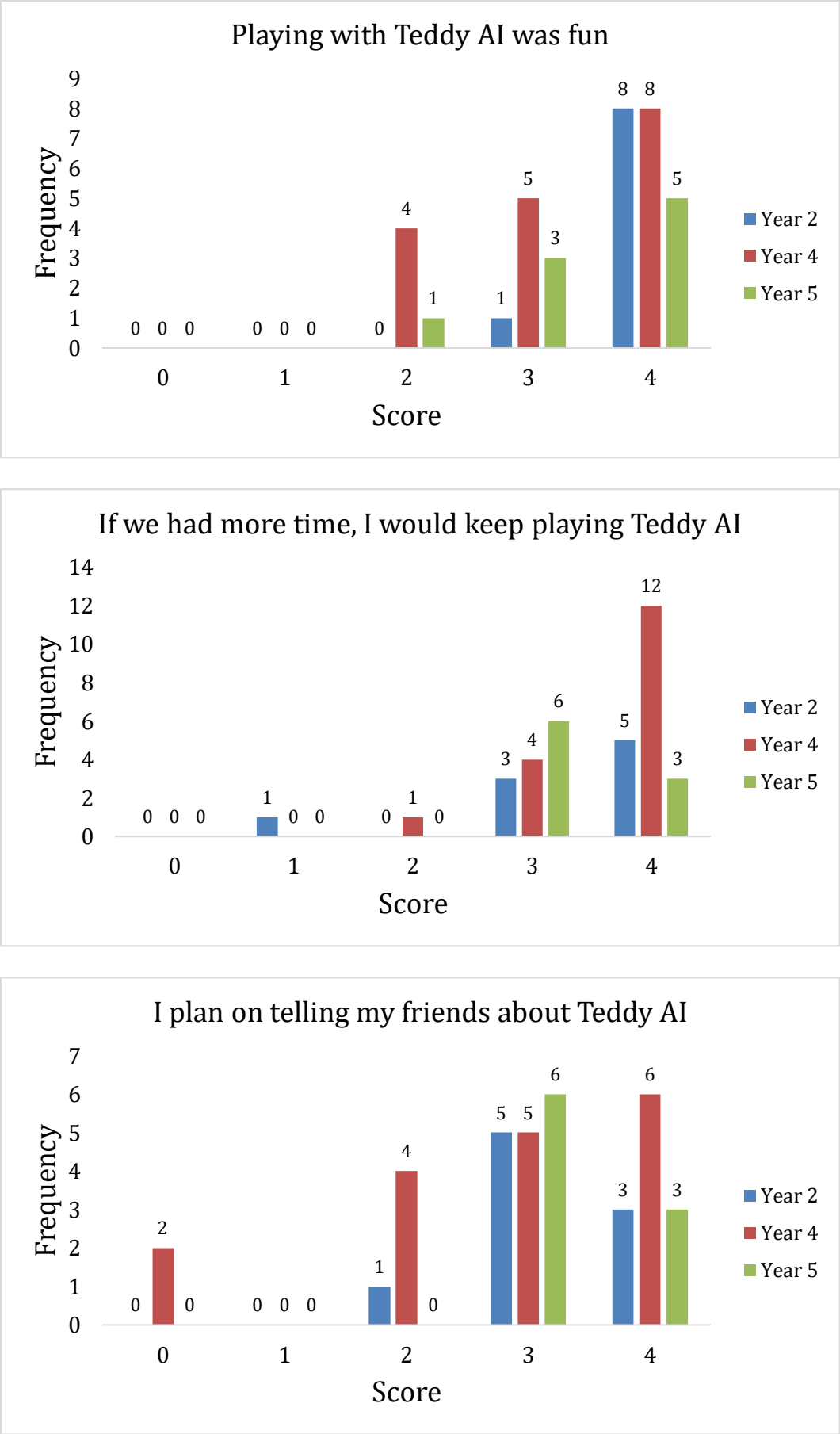

Figure 3. Results of additional questions split by year group.

4 CONCLUSIONS

The primary objective of this study was to investigate the usability, user experience, and potential educational benefits of an AI conversational learning application, Teddy AI, among primary school students with an average age of 8 years. Our study on the Teddy AI app has revealed interesting implications for the development of AI-enabled educational tools. The ease of use, user engagement, and satisfaction experienced by children were key points of our analysis. Using a combination of gamified exploration, and open-ended follow-up questions, the study aimed to highlight critical aspects of the application and gather user feedback for improvement. As previous research [10] has consistently shown that game-based learning has the potential to boost motivation and promote productive and meaningful learning experiences which corresponds with the observation in our study. The user-friendliness of the interface, ease of navigation, and interactive display ensure a smooth and enjoyable learning process for children. Needless to say, Teddy AI scores high on this parameter. The various features, such as the lessons, map, and quizzes, provided a diverse and interactive learning environment, fostering greater user engagement. Additionally, the reward system, which included earning coins and verbal appraisal, also served as positive reinforcement, motivating the students to continue using the application and improving their learning experiences. The use of the SUS provided valuable insight into the participant's subjective experience of the app and indicated a generally positive experience with a few areas for improvement. Based on the average SUS scores, the app's usability and reliability were rated highly by participants. Additional Likert-type questions and open-ended follow-up questions also allowed for a more in-depth understanding of the participants' satisfaction and preferences and helped identify areas that could be improved in future iterations of the app. The Teddy AI App's ability to record and analyse user interactions allowed for personalised feedback and adaptation, further enhancing the learning experience. For instance, the application could identify when a user struggled to recognise or pronounce a word correctly and provide motivation and guidance to improve the user's performance [11]. Though the findings indicated significant satisfaction among children, the study also shed light on some areas needing refinement. The detailed analysis highlighted the necessity for significant customisation based on the students' individual needs. Several children faced difficulties in pronunciation recognition, and some found understanding the quizzes challenging due to their age. Further refinements and customisation based on users' language proficiency and comprehension levels could enhance learning outcomes and user experience. Similar principles have been applied in other intelligent tutoring systems and chat-oriented dialogue systems to provide personalized feedback and adapt to the user's needs [12].

Despite the overall positive implications, some factors should be under consideration while analysing the research outcomes. The sample size of 35 participants may not provide a comprehensive representation of the population at large, and this study can be considered as a preliminary investigation into the application's usability and effectiveness. Additionally, while the screen recordings provided valuable insights for researchers, different devices used during gameplay may have affected the data collected. Lastly, the study focused on a specific version of the application, and the findings may not be generalizable to other versions or similar applications. Future areas of research could include expanding the sample size to better represent the target demographic and examining other aspects of the application, such as comparing the effectiveness of the conversational AI approach to other learning methods. Additionally, a longitudinal study could help assess the improvement in students' learning outcomes over time and help develop more effective AI applications for educational purposes. Moreover, personalisation is an important feature in the Teddy AI app, which has proved to be an important factor in achieving learning objectives. However, this could be augmented with further customisations based on individual learning speed, comprehension, and attention span, creating room for more inclusivity.

To conclude, this study provides valuable insights into the usability and user experience of Teddy AI, an AI conversational learning application aimed at younger children. The findings suggest that this type of application can effectively promote learning and engagement among primary school students, with the potential for further optimization and growth in the rapidly evolving domain of AI-assisted education. The Teddy AI app had a positive overall user experience. However, incorporating enhancements in personalization and refining certain areas for instance better microphone input capacity and diversified visual contents could significantly improve the experience. The findings of the user study could serve as a critical resource to educators and developers while designing, implementing, and refining AI-based educational tools. Future studies could consider a larger pool of participants and incorporate enhancements based on the outcomes of the study.

ACKNOWLEDGEMENTS

We would like to thank the participating schools for their support with conducting the study and the children for taking part in the study. We would also like to thank Ms R. Thavagnanaratnam for their support with the data collection.

REFERENCES

[1] Snyder, Lisa Gueldenzoph, and Mark J. Snyder. "Teaching Critical Thinking and Problem Solving Skills." Delta Pi Epsilon Journal, vol. 50, no. 2, 2008.

[2] Wrenn, Jan, and Bruce Wrenn. "Enhancing Learning by Integrating Theory and Practice." International Journal of Teaching and Learning in Higher Education, vol. 21, no. 2, 2009, pp. 258-265.

[3] Woolf, B. P. (2010). Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning. Morgan Kaufmann.

[4] Johnson, W. L., Lester, J., & Spires, H. (2014). Artificial intelligence in support of computational models of engagement in narrative-centered learning environments. In R. Sottilare, A. Graesser, X. Hu, & B. Goldberg (Eds.), Design Recommendations for Intelligent Tutoring Systems: Adaptive Learning Experiences (Vol. 2, pp. 173-181). U.S. Army Research Laboratory.

[5] Khosravi, Hassan, et al. "Explainable Artificial Intelligence in Education." Computers and Education: Artificial Intelligence, 2022.

[6] Ige, O.A., Hlalele, D.J. Effects of computer-aided and blended teaching strategies on students' achievement in civic education concepts in mountain learning ecologies. Educ Inf Technol 22, 2693–2709 (2017). https://doi.org/10.1007/s10639-017-9598-x

[7] Sauro, J., & Lewis, J. R. (2016). Quantifying the User Experience: Practical Statistics for User Research - 2nd Edition. San Francisco, CA: Morgan Kaufmann.

[8] Rohles, Björn, et al. "Creating Positive Learning Experiences with Technology: A Field Study on the Effects of User Experience for Digital Concept Mapping." Heliyon, vol. , 2022, Elsevier.

[9] Lindgaard, G., & Dudek, C. (2003). What is this evasive beast we call user satisfaction? Interacting with Computers, 15(3), 429-452.

[10] Jääskä, Elina & Lehtinen, Jere & Kujala, Jaakko & Kauppila, Osmo. Game-based learning and students' motivation in project management education. Project Leadership and Society. 3. 100055. 10.1016/j.plas.2022.100055, 2022.

[11] Graesser, A. C., Chipman, P., Haynes, B. C., & Olney, A. (2005). AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. IEEE Transactions on Education, 48(4), 612-618. DOI: 10.1109/TE.2005.852595

[12] Banchs, R. E., & Li, H. (2012). IRIS: A chat-oriented dialogue system based on the vector space model. Proceedings of the ACL 2012 System Demonstrations, 37-42. DOI: 10.5555/2390477.2390486 URL.