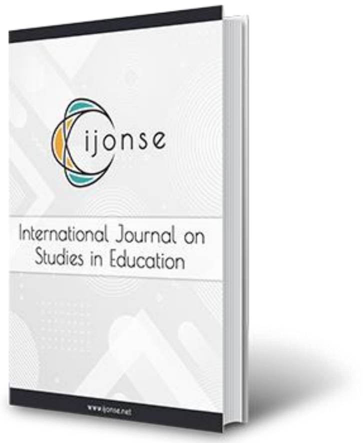

www.ijonse.net

Enhanced Educational Experiences through Personalized and AI-based Learning

Vivian Kretzschmar
Stuttgart Media University, Germany

Alica Sailer
Stuttgart Media University, Germany

Michael Wertenauer
Stuttgart Media University, Germany

Jürgen Seitz
Stuttgart Media University, Germany

To cite this article:

Kretzschmar, V., Sailer, A., Wertenauer, M., & Seitz, J. (2024). Enhanced educational experiences through personalized and AI-based learning. *International Journal on Studies in Education (IJonSE), 6*(2), 191-209. https://doi.org/10.46328/ijonse.206

International Journal on Studies in Education

2024, Vol. 6, No. 2, 191-209 https://doi.org/10.46328/ijonse.206

Enhanced Educational Experiences through Personalized and AI-based Learning

Vivian Kretzschmar, Alica Sailer, Michael Wertenauer, Jürgen Seitz

Article Info

Article History

Received:
13 October 2023
Accepted:
26 February 2024

Keywords

AI
AI-based learning
AI education
Video-based learning
Self-learning

Abstract

Academic motivation is a pivotal factor in shaping students' educational trajectories. Mathematics education holds a unique position as a school subject, due to its influence on students' grade transitions and cognitive development at both basic and advanced levels. From March to May 2023, the AI Education (AIEDN) research project investigated how an AI-based learning assistant can enhance students' understanding of the subject matter through video-based learning. For this study, 275 students were selected in the age range of 14-20 from two secondary (N=137) and two grammar schools (N=138) in Baden-Württemberg, Germany. The quantitative experiment tested the extent to which learners solve more tasks, build broader (transfer) knowledge, and retain it. Students were given a set of mathematical problems on an unknown topic to solve in 90 minutes. The test group used the AI assistant to ask questions, while the control group used only keyword searches. The results of the t-test indicated that the test group who used the AI Learning assistant in the advanced course achieved a significantly strong (Cohen's $d=.63$) increase in performance results ($T(19)=-2.82$; $p<.01$). Conversely, the advanced course control group as well as both the test and control groups of the basic course(s) showed no significant improvements.

Introduction

Today, information technologies such as e-learning tools and Massive Open Online Courses play a central role in education, especially in terms of transforming and simplifying access to knowledge and learning (Palacios Hidalgo et al., 2020). It is increasingly emphasized that Artificial Intelligence (AI) should not be seen as a mere tool to support the educational landscape, but that its key role can permanently change the way we learn and teach collaboratively.

AI can be defined as a scientific discipline of computer science, as a collective term for techniques used to enhance the performance of hardware and software programs, as a hardware and software system that exhibits intelligent problem-solving behaviour, or as an artificial being with intelligence (Lämmel & Cleve, 2020). In this publication, artificial intelligence is understood as an intelligent system.

Researchers highlight the ability of AI to personalize education through a higher level of interaction and individualized feedback. In this context, the concept of adaptive learning, in which the method of knowledge transfer adapts individually to the learner's environment, e.g. in the form of learning preferences or existing knowledge, is gaining importance. The processes in our memory that control learning are adaptive if they prioritize meaningful choices and produce positive results. Therefore, this interplay between adaptive learning and different memory processes promotes utility-oriented behaviour at every stage of life (Ahmad et al., 2022; Ciolacu et al., 2018; Hartley et al., 2021; Ouyang et al., 2022; Popenici & Kerr, 2017; Schön et al., 2023).

This approach not only increases learners' motivation and interest, but also allows them to actively shape their educational journey, which can lead to improved learning performance (Maedche et al., 2019; Schön et al., 2023). Not only with the development of digital assistants by Big Tech (Maedche et al., 2019), but also with the recent successful launch of Chat-GPT 3 on November 30, 2022, the use of artificial intelligence (AI) as another everyday tool has arrived and been accepted by the general population (Kim et al., 2020; Lee & Perret, 2022). However, examples of its use in academic settings are still rare (Alam & Mohanty, 2022). In this context, learning with videos is becoming increasingly popular (Engels & Schüler, 2020; Kokoç et al., 2020). However, individual knowledge gaps can only be closed to a limited extent due to the prevailing oversupply of different platforms (Vonderau, 2016). Intelligent tutoring systems (ITS) can already support students' learning with tailored solutions and thus increase the effectiveness of learning (Liu et al., 2021).

An ITS enables a customized tutorial solution. This allows students to learn individually at their own pace and increases classroom effectiveness (Liu et al., 2021). Users decide for themselves when to learn and what 'customized content' to engage with (Chen et al., 2020). In this way, learning barriers are reduced and processes optimized (Ahmad et al., 2022). Simultaneously, AI provides immediate (audio-visual) feedback to its users (Schleiss & Göllner, 2022). This enables faster, interactive, and personalized knowledge acquisition. Learning in a self-defined learning atmosphere serves to promote long-term retention of information (Chen et al., 2020). Since each student has different abilities and levels of education, it can be challenging for teachers to meet individual demands (Ahmad et al., 2022). It is therefore important to balance high-performing students and lower-performing students (Chassignol et al., 2018). The use of an ITS is intended to close this gap and to counteract the ongoing teacher shortage (BMBF, 2015).

Method

The AI Education (AIEDN) study involved 275 students from two different school types and investigated how a semantic AI assistant could compensate for existing weaknesses in a more targeted way and enable and reinforce better (mathematical) understanding in the learning process through video learning. Preliminary results show that learning with explanatory videos leads to positive learning gains and enables longer-term recall of procedural knowledge (Lloyd & Robertson, 2012; Van Der Meij & Van Der Meij, 2014).

The quantitative experiment was conducted from March to May 2023. Students of two different standards (9th and 11th) and school types (grammar vs. secondary school) were given a fixed set of mathematical problems to solve on a previously unknown topic, familiarizing themselves with these either with the AIEDN AI assistant, or with the help of an identical simulation, operating through a regular keyword search. To check the learning effect, the test was repeated after 6-8 days with similar tasks of the same scope without an assistant. In summary, it can be said that the AIEDN AI assistant led to significant improvements in learning performance, especially for the target group of students in the upper performance range.

Research Hypotheses

Video-learning focuses on the comprehensibility of the instructional content, which is made possible by the interaction between the presenter and the recipient (Findeisen et al., 2019). AIEDN allows for a more intensive engagement with (mathematical) topics and at the same time a more individualized approach to learners through the selection of suitable content. While it does not yet provide a fully personalized experience, the assistant contributes to a significant level of individualization for the user. As a result, it can be expected that the AI assistant will facilitate more positive learning effects, initiating a higher efficiency than traditional video learning.

This leads to three hypotheses, which have been tested in a mixed-methods study:

1. Learners can complete more tasks at the same time with the help of the AI learning assistant or when they have learned with the AI tool, than learners who do not have the AI assistant available, or who have learned without it.
2. Learners who used the AI learning assistant can build deeper knowledge and apply it more effectively to transfer tasks than learners who did not use the AI learning assistant.
3. Learners who learned with the AI learning assistant can recall the acquired knowledge for longer time periods than those who learned without AI assistance.

Research Design

Sample

Within this empirical study, 275 students from four schools in Baden-Württemberg, Germany, were examined constituting two secondary schools and two grammar schools, whereby both school types are to be considered as separate experiments. This allows the AI assistant to be tested on a heterogeneous group of students. The experiments at the two school types were conducted separately, with 139 students at the grammar schools and 137 students at the secondary schools. Careful consideration was given to ensure an overall gender distribution. The selection of schools was limited to Baden-Württemberg to exclude differences in curricula between the individual federal states, thereby ensuring comparability. The students from grammar schools (grade 11) and secondary schools (grade 9) were set to graduate soon to ensure equal or higher participatory motivation due to the possibility of thematic repetition of learning material for the final exams. By focusing on one grade level per experiment, age-related differences in maturity and learning development were eliminated.

Study Setup

To empirically test the research hypotheses, an experiment was conducted from March to May 2023. Students of both grade levels were given a fixed set of seven (for secondary schools) or eight (for grammar schools) mathematical problems on a previously unknown topic within a 90-minute time window (based on the scope of a regular exam situation). Depending on the test group, they were able to familiarize themselves with this topic either with the AIEDN AI assistant or with the help of a regular keyword search on a "simulated AIEDN video platform". Assignment to the test or control group was randomized. The mathematical tasks did not differ, to ensure comparability between the two groups.

Mathematics was chosen because of its objective evaluation and comparison possibilities. To check the learning effect, the test was repeated after 6 to 8 days with similar tasks of the same scope without the use of the AI assistant. All questions were answered in the same way as they would be in a normal school day, i.e. on paper, to ensure constant visibility of all questions at all times and to allow for a flexible processing sequence.

Pretest

The preliminary research served as the basis for the elaboration of the study design and the development of the first prototype of the AI assistant. To validate the planned study design, a pretest was conducted with 22 grammar school students in 9th grade. Since the pretest was conducted at a grammar school that was also a regular participant in the study, the 9th grade was selected using the task set of the secondary school, so that an exchange between the students could not be of help to the later test subjects. The participants were not aware at the time of participation that this was a pretest. As a result, the task set for the study was extended by two tasks to ensure that no student would be able to complete the test within the allotted time.

Screening and Questionnaire

After completing the test on the first day of the study, a screening questionnaire was administered to assess subjects' learning history, demographic characteristics, migration background, and previous experience with educational videos. A semi-standardized questionnaire (Likert scale) consisting of open and closed questions was used to evaluate the participants' learning experience with the learning assistant after the test. Both the screening and the questionnaire were linked to the answer sheet of the math test by a code for anonymous evaluation. In addition, the learning experience of randomly selected students was recorded through qualitative interviews after the test in order to gain a deeper insight into the level of understanding and motivation of the learning process that took place.

Structure & Functionality of the AIEDN AI Learning Assistant

As part of this study, a prototype based on Streamlit, was implemented (Figure 1). Streamlit is a Python-based library that enables the display of text, interaction elements, and output of AI models, as well as the visualization of data and model performance, combined with the ability to adjust model input parameters via the user interface. The use of Streamlit enables the rapid creation of simple web applications for AI models.

The prototype was functionally constrained to meet the requirements of the study design. This was necessary to draw appropriate conclusions and to ensure comparability between the test and the control group within the study. To avoid bias, no adjustments were made to the prototype during data collection. To ensure comparability within the study, both the AI-based and non-AI-based systems were developed and deployed with the same graphical user interface. The data base of the prototype consisted exclusively of explanatory math videos by Daniel Jung. The recordings of these videos were transcribed and divided into segmented sections. After, the resulting transcripts were transferred to the semantic AI environment. For each text segment, the corresponding video, the associated timestamp, and additional metadata such as topic area and YouTube playlist were stored. A semantic fingerprint for each sentence of the transcript(s) enables the AI to understand the content of the video on a semantic level, generating relevant answers to queries.

The semantic AI semantha® enables meaning-based search in video content. The effectiveness of the semantic search function increases in proportion to the amount of context provided for the question asked. Participants in the AI group were instructed to ask the AI-based learning assistant questions with high precision, analogous to interacting with a human tutor. Participants in the control group, on the other hand, were asked to enter keywords to get the most relevant results.

Figure 1. User Interface of the AIEDN Streamlit Prototype

Results

To investigate a significant increase in performance with the use of the AI learning assistant, several t-tests were carried out to compare the success in mathematical problem-solving either with the AIEDN AI learning assistant or the optically identical simulation as a function of their mathematical advancement (basic and advanced course), the average number of videos viewed during task processing, and the overall frequency of use of the learning videos.

Analysis by Basic and Advanced Course

In a subsequent investigation of the differences between the basic and advanced course levels in grammar schools (Figure 2 & 3), it was shown that the AI assistant benefitted especially the already high-achieving students. The results of the t-test also indicate that within the group of subjects who used the AI assistant in the advanced course, there is a significant change in the results measured by the mean of the average scores achieved between the first assessment ($M1=32.6$; $SD1=15.84$) and the second one ($M2=42.95$; $SD2=25.78$) ($T(19)=-2.82$; $p<.01$). Thus, it can be said that there is a strong effect (d=.63) in terms of improvement in scores due to the use of the AI assistant.

In contrast, the control group without AI in the advanced course showed no significant improvement in performance between the two study days (M1=29.27; SD1=20.67; M2=29.19; SD2=22.86; T(30)=0.02; p>.05; d=.004). Within the test group of basic courses that used the AI assistant, an increase in scores between the two survey days was observed by evaluating the means (M1=10.87, SD1=10.33; M2=13.72; SD2=17.31), but no significant improvement in performance occurred (T(45)=-1.35); p>.05) despite a small effect size (d=.20).

The same applies to the control group in basic courses. The analysis of the mean scores on both study days revealed a minimal increase (M1=7.83; SD1=8.08; M2=8.01; SD2=11.02), but no significant change in the performance spectrum (T(40=-.23; p>.05; d=.04). Thus, a positive correlation could be established between the hypotheses regarding the enrollment in advanced courses or basic courses, with only the students in advanced courses, but not those in basic courses, benefiting significantly from the use of the AIEDN AI assistant.

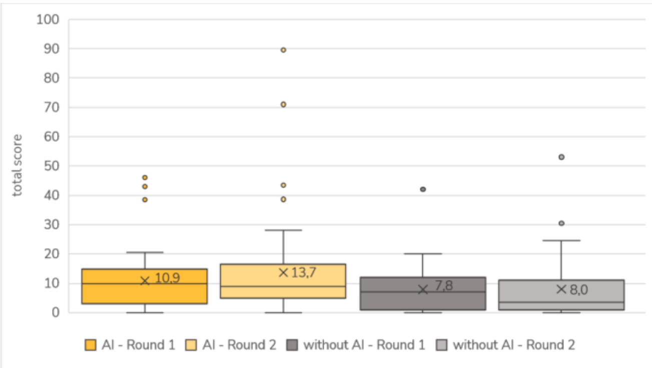

Figure 2. Basic Level Mathematics Courses at Grammar Schools

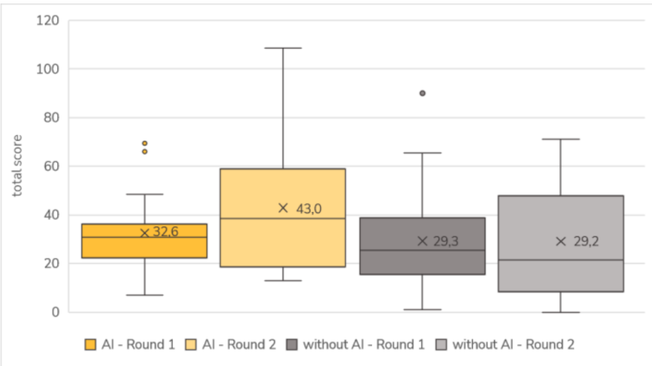

Figure 3. Advanced Level Mathematics Courses at Grammar Schools

Analysis by the Average Number of Videos viewed during Task Processing

To assess a possible increase in learning effect depending on the frequency of use of AIEDN learning videos, participants were asked at the end of the study to self-assess how often (1 to 10 times) they used the learning videos to work on an average task using a Likert-type scale (Figure(s) 4,5 & 6).

Within the group of grammar schools that used an average of 1-3 AIEDN videos to complete tasks, there was a significant increase in total score when looking at the mean scores between the two survey points (M1=19.54; SD1=16.91; M2=24.22; SD2=26.38). The corresponding t-test evaluation also showed significance of the test group's performance improvement with a medium effect size (T(46)=-2.14; p<.05; d=.31).

The test group that took an average of 4 or more videos to solve the math problems, showed an increase in performance between study day 1 (M1=9.13; SD1=8.87) and study day 2 (M2=16.69; SD2=18.01) that exceeded the overall average of the average scores achieved. In addition, this group of students also showed an overall significant increase in learning or performance with a medium effect size (T(15)=-1.91; p<.05; d=.48).

In contrast, the control groups of grammar schools that received the visually identical non-AI-based system showed a minimal increase in their mean total score (M1=20.89; SD1=20.22; M2=21.76; SD2=22.82) within the subgroup that used 1-3 videos to complete tasks, but did not achieve a significant, effect size improvement in their overall learning performance (T(41)=-0.32; p>. 05; d=.05).

Similar phenomena were observed in the subgroup of grammar schools without AI, which played significantly more content to solve the assigned tasks, with >4 videos. However, in contrast to the previous group, this resulted in a deterioration of the total score achieved above the mean between the two survey points (M1=10.11; SD1=9.85; M2=9.59; SD2=11.46). Furthermore, identical to the subgroup without AI who used 1-3 videos, no significant improvement in students' learning performance could be evaluated (T(27)=0.41; p>.05; d=.08).

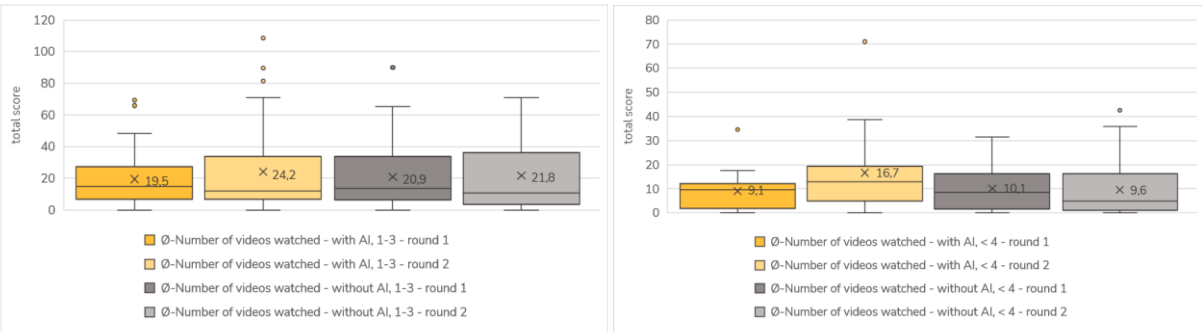

Figure 4. Grammar Schools with an Average of 1-10 Learning Videos played

The secondary school study group benefited from the use of the AI assistant to varying degrees, depending on how intensively the learning assistant was used. Both subgroups that used either 1-3 (M1=5.06; SD1=5.39; M2=7.66; SD2=12.87) or >4 videos (M1=3.25; SD1=3.70; M2=4.58; SD2=6.86) to solve the assigned tasks were able to improve their overall scores on average over both study days. While the group of participants who played significantly more videos for support with >4 videos was able to achieve a significant increase in their learning outcomes with a small effect size (d=.26) (T(52)=-1.92; p<.05), no such increase was measurable in the subgroup with AI who used less content with 1-3 videos played (T(15)=-0.84; p>.05; d=.21).

The secondary school groups that were not provided with an AI assistant provided the same pattern of results in terms of the number of videos watched associated with a demonstrable increase in performance. Again, it was observed that students who tended to require >4 videos as support showed both an increase in performance on their overall score during both study days (M1=3.74; SD1=5.03; M2=6.01; SD2=8.91) and a significant correlation associated with the intended learning improvement through AIEDN with a medium effect size (T(38)=-1.99; p<.05; d=.32). In contrast, participants who used less than 4 videos as support materials showed a decrease in total score at both time points (M1=5.31; SD1=5.59; M2=4.60; SD2=7.38) and a nonsignificant improvement in overall performance (T(23)=0.41; p>.05; d=.08).

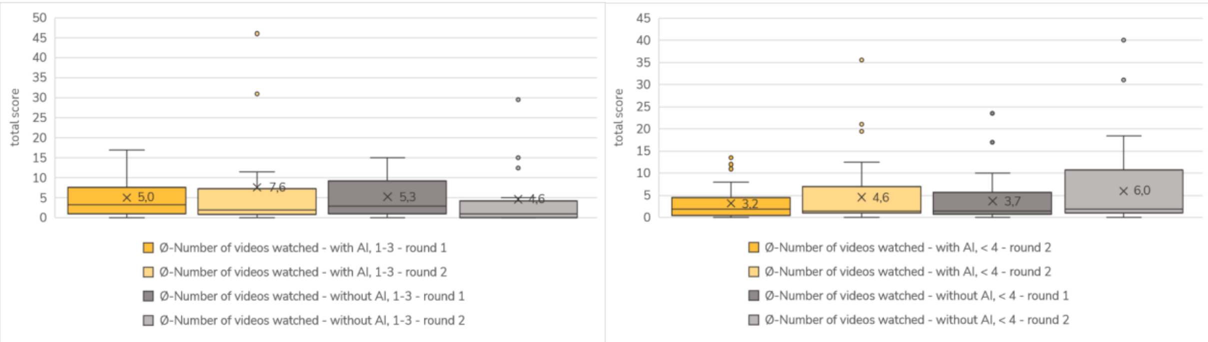

Figure 5. Secondary Schools with an Average of 1-10 Learning Videos played

Across the study, regardless of school type, students using the AIEDN AI assistant improved their overall scores on both study days. Nuanced differences emerged: the group that needed fewer videos (1-3) to complete the tasks achieved significantly higher overall scores on average (M1=15.87; SD1=16.11; M2=20.02; SD2=24.69) than the group with AI that watched an average of 4 or more videos (M1=4.61; SD1=5.84; M2=7.39; SD2=11.58). However, both the 1-3 video sample group (T(62)=-2.30; p<.05) and the >4 video sample group (T(68)=-2.55; p<.01) show an overall significant improvement in performance with small (d=.29) to medium (d=.31) effect sizes.

Unfortunately, unlike the previous observations, these cannot be confirmed for either subgroup without the use of AI. Although both the group with 1-3 (M1=15.23; SD1=18.05; M2=15.52; SD2=20.42) and the group with >4 used learning videos (M1=6.40; SD1=8.01; M2=7.51; SD2=10.13) achieved a minimal increase in their average total score, neither the use of a smaller number of learning videos (T(65)=-0.16; p>. 05; d=.02) nor a higher number (T(66)=-1.29; p>.05; d=.16) produced a significant improvement in learning performance without the use of AI.

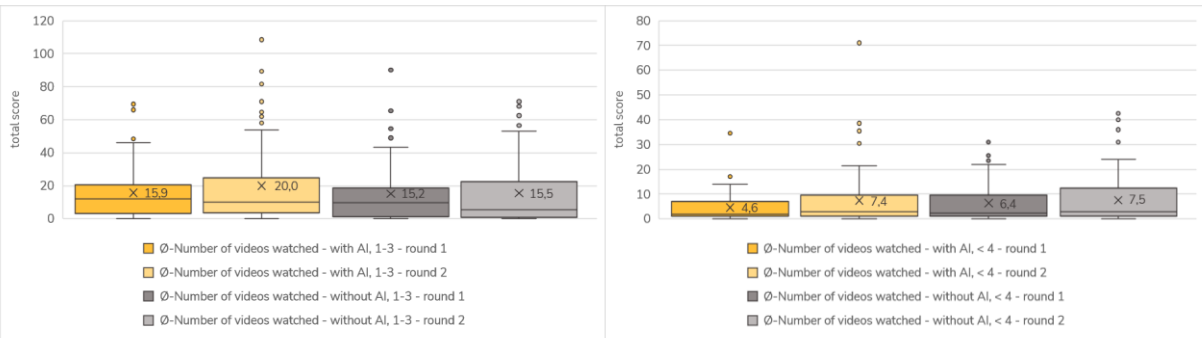

Figure 6. Total Participants with an Average of 1-10 Learning Videos played

Analysis by Frequency of Use of the Learning Videos

In a further investigation, a self-assessment by the participants was used to determine whether the use of explanatory or instructional videos in the run-up to the study led to an improvement in learning. For this purpose, a Likert scale from 1 (always) to 7 (never) was used, whereby the value 4 as an 'escape category' was not included in the evaluation (Figure(s) 7,8,9,10,11 & 12).

In the grammar school students' group, both students who had previously used learning or instructional videos frequently (1-3) and those who had used them less frequently or never (5-7) were able to increase their overall mean score by using the AI assistant. It is noticeable that the mean value of the achieved total score of test day 1 and test day 2 and the increase in performance of the students who had rarely or never been in contact with learning videos before was significantly higher (M1=22.73; SD1=17.13; M2=28.41; SD2=28.45) than that of the subjects who used them (very) frequently (M1=9.26; SD1=7.31; M2=15.8; SD2=17.36).

However, the analysis of the corresponding t-tests clearly showed that only the group that used AIEDN and had previously studied frequently (1-3) with learning videos showed a significant learning improvement with a medium effect size (T(24)=-2.36; p<.05, d=.47). Students who also tested AIEDN for the first time during the study but had little to no prior experience with online learning (5-7) did not experience a significant change in their learning performance, despite a medium effect size (T(27=-1.68; p>.05; d=.32).

It could be shown that the group of grammar school students who had solid prior knowledge (1-3) in using instructional and explanatory videos, but who did not receive an AI assistant for solving the mathematical problems, showed neither an increase in their overall mean scores between the two study days (M1=10.44; SD1=8.87; M2=8.76; SD2=11.16), nor an overall significant improvement in their learning performance (T(26)=1.05; p>.05; d=.20). In contrast, the subgroup that also did not use AIEDN during the study but received the identical prototype without AI and also had little to no experience with instructional videos, again showed significantly higher and above average overall scores (M1=23.57; SD1=22.50; M2=24.54; SD2=23.74). Despite this increase, no overall significance of change or improvement in learning was found (T(34)=-0.31; p>.05; d=.05).

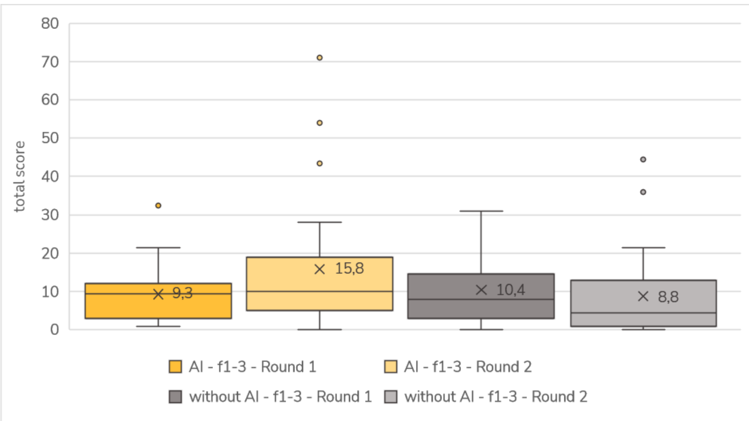

Figure 7. Grammar Schools with Frequent (1-3) Prior Use of Learning and Explanatory Videos

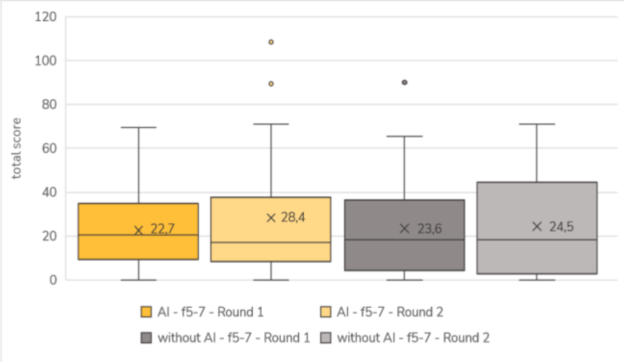

Figure 8. Grammar Schools with Less Frequent (5-7) Prior Use of Learning and Explanatory Videos

Looking at the secondary school students who were part of the test group, it is noticeable that, contrary to all previous evaluations, on average their total score deteriorated from study day 1 to study day 2, despite intensive prior use of learning or explanatory videos as well as additional support from the AIEDN AI assistant (M1=2.11; SD1=3.83; M2=1.93; SD2=2.89). The analysis of the t-test also showed that no significance was found in terms of improvement in their learning performance (T(13)=0.17; p>.05; d=.04).

Groups of secondary school students who also had access to the AI assistant, but who had little or no experience with learning or instructional videos, were able to both increase their overall score between the two runs (M1=3.86; SD1=4.20; M2=6.17; SD2=9.89) and achieve a significant change or increase in their overall learning performance (T(47)=-1.92; p<.05; d=.28).

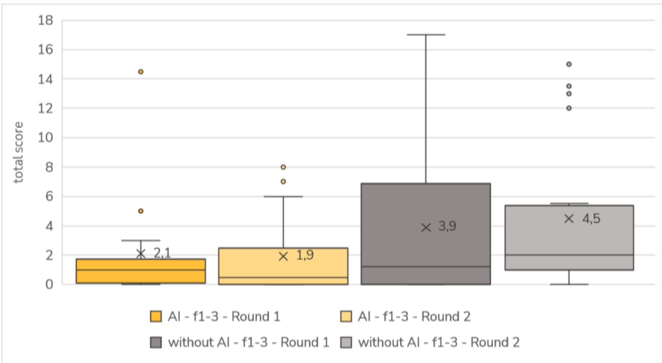

Figure 9. Secondary Schools with Frequent (1-3) Prior Use of Learning and Explanatory Videos

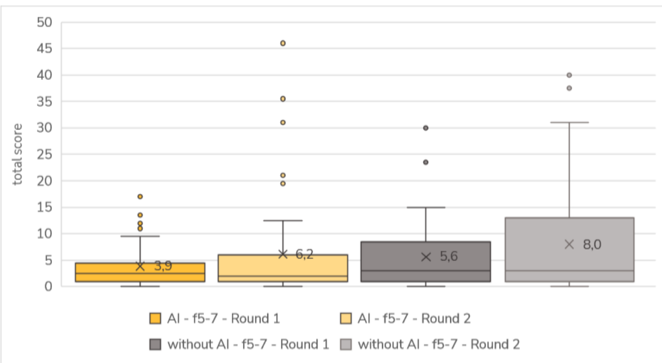

Figure 10. Secondary Schools with Less Frequent (5-7) Prior Use of Learning and Explanatory Videos

The control groups of the secondary schools were able to achieve an average increase in their total score in relation to the learning and explanatory videos, regardless of their prior knowledge, both with high (1-3) (M1=3.89; SD1=5.01; M2=4.5; SD2=5.14) and with low (5-7) (M1=5.61; SD1=6.97; M2=8.00; SD2=11.06) prior knowledge. It was clear that the group with lower prior knowledge (5-7) again achieved significantly higher total scores. However, it was also found that between the two groups, neither students with high (T(17)=-0.46; p>.05; d=.11) nor low prior knowledge (T(36)=-1.61; p>.05; d=.26) were able to achieve a significant change in their learning performance without using the AI assistant.

Overall, the evaluation of the total participants, regardless of the type of school, who used AI to solve the task shows that both students with high (M1=6.69; SD1=7.13; M2=10.82; SD2=15.46) and with low reference and prior knowledge (M1=10.82; SD1=14.16; M2=14.36; SD2=22.28) achieve a higher total score regarding the use of instructional and explanatory videos in everyday school life. What is striking here, as in other analyses, is that students with lower prior use of instructional and explanatory videos achieve higher performance or scores, regardless of whether they belong to the test or control group. In each case, the use of AI contributes to a significant improvement in performance, regardless of prior knowledge of relevant tools, for both high (T(38)=-2.20; p<.05, d=.35) and low (T(75)=-2.44; p<.01; d=.28) baseline knowledge subjects.

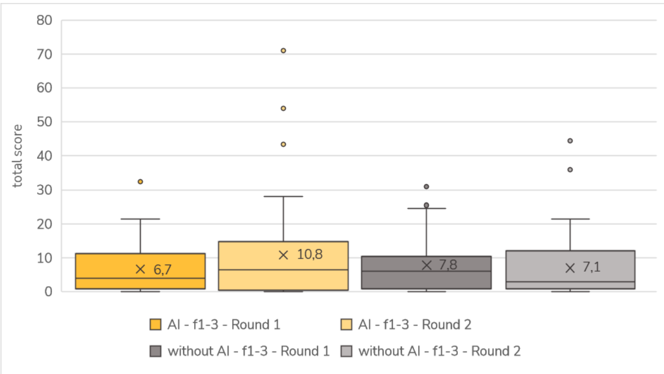

Figure 11. Total Participants with Frequent (1-3) Prior Use of Learning and Explanatory Videos

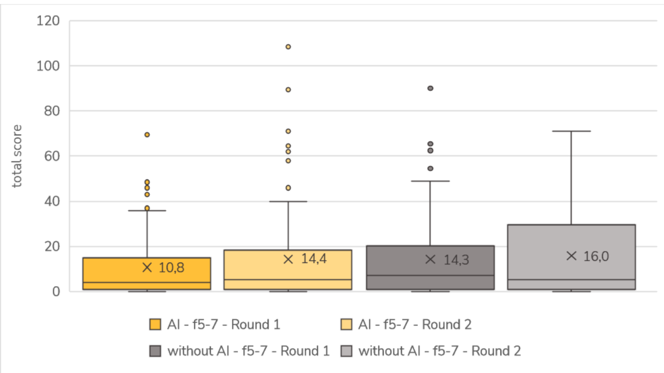

Figure 12. Total Participants with Less Frequent (5-7) Prior Use of Learning and Explanatory Videos

Participants who did not have the AI available to work on the math problems presented, performed worse overall in the subgroup that used explanatory videos more often for learning, minimizing them between the two runs

(M1=7.82; SD1=8.17; M2=7.06; SD2=9.40). In contrast, the low prior use group increased their total score over the test period (M1=14.34; SD1=18.68; M2=16.04; SD2=20.03). However, neither the high prior use control group (T(44)=-0.70; p>.05; d=.10) nor the low prior use control group (T(71)=-1.01; p>.05; d=.12) found significance in the improvement of learning performance without the use of AI.

Discussion

The results of the study shed a differentiated light on the effects of using the AI assistant between students at the basic and advanced course levels. In the advanced math course, the test group using AIEDN steadily increased its overall score, while the control group without AI deteriorated. A closer look at the test group revealed that the use of AI led to significant learning improvements, while the control group did not show comparable changes or improvements. Possible reasons for this could be seen in an excessive demand of the additional tool, which requires a certain acclimatization period for some students, since especially advanced course students are trained to develop a more detailed theoretical-scientific focus than basic course students, who are limited to a basic level of understanding (Staatsministerium für Kultus, 2016).

In the basic courses, both the students who used the AI and those who received only the identical prototype without the AI showed an increase in overall scores over the entire study period. However, there was no change or improvement in the range of performance within the groups. These results suggest that the use of AI assistants may be particularly beneficial in already high-performing courses such as advanced mathematics. In addition, it can be assumed that AIEDN favors "subject-affine" students in a special way, since knowledge is conveyed more intensively visually. Here, students can benefit from the adaptive learning features of AI to deepen their knowledge and achieve their learning goals more effectively. In contrast, in basic courses, where there may be a wider range of performance, the use of AI alone does not appear to be sufficient to significantly improve overall performance levels. These findings underscore the importance of purposefully integrating AI technologies into education and show that the effectiveness of such tools is highly dependent on the individual needs, prerequisites, and challenges of students.

The grammar school students who worked with AI support were able to achieve a significant improvement in performance, regardless of how many videos they had watched on average per task they worked on. In contrast, the control group did not show any significant improvement. In fact, those who watched four or more videos performed worse. The secondary school students who used AIEDN and watched an average of one to three videos improved their performance, but not significantly. Students who watched four or more videos, however, did. Secondary students without AI support who watched one to three videos worsened, while those who watched four or more videos improved significantly.

Overall, students who watched an average of one to three videos scored higher than those who watched four or more videos. There was a significant improvement in scores for the students with AI support compared to the group without AI support. Since this is also true for the scores of all participants, the question arises as to what extent the average number of videos watched per task is related to the results. The higher average score of those who watched only one to three videos per task suggests that the higher performing students had to watch fewer videos.

The grammar school students who worked with AI support and who reported frequent use of learning videos, a significant improvement in learning was demonstrated. In contrast, students who had less experience with learning videos did not show significant improvement. However, they scored higher on average. The students from secondary schools in the AI-assisted group, regardless of their previous experience with learning videos, did not show any significant improvement, and the group with frequent use of learning videos worsened in terms of scores. The secondary students in the control group showed no significant improvement in performance, with the group with less video experience scoring higher overall. Altogether, it can be said that those who use instructional videos less frequently achieve higher scores on average. Again, it can be hypothesized that lower performing students have a greater need for supplemental materials such as explanatory videos and therefore use them more frequently. There was a significant increase in performance in the group of students with a lot of experience with learning videos, making them the beneficiaries of the tool.

The analysis of the interviews conducted after the study days shows that the AI assistant was well received by most of the students. Interestingly, many of them had no prior knowledge in the field of artificial intelligence but found the idea of being assisted by such technology motivating. Even participants in the control group, who had to use a visually identical simulation, referred to "the AI" as having helped them.

Within the interviews, students frequently emphasized the importance of personalized support from the AI. They wanted the AI to tailor their answers to their individual background and prior knowledge. In this way, weaker students would receive better support.

Conclusion

To investigate the impact of a semantic AI assistant (AIEDN) on student learning outcomes, an extensive mix-methods research study was conducted at two grammar schools and two secondary schools in Baden-Württemberg, Germany. A prototype was developed to optimize video searches by semantically understanding search queries and providing relevant answers in text form, while also referring to the relevant sections in the mathematical videos by YouTuber Daniel Jung, which the AI assistant is trained with. Due to the simplified interactions and more efficient, "filtered" query processes compared to traditional keyword searches, it was expected that the AI assistant could help students learn more effectively during self-study phases, solve a greater number of problems, and retain knowledge for a longer period using topic-related, explanatory, educational video content. In addition, the study investigated whether the use of the learning assistant could contribute to an overall higher learning motivation.

After analyzing the data, the results showed that students who used the AI assistant were able to apply what they had learned better than the control group, which had access only to traditional keyword searches. The results related to all three examined categories generally support Hypothesis 1. Across all participant groups, those who used the AI learning assistant achieved a significant increase in their mean total scores between study days, indicating improved task completion efficiency.

The effect size was strong (d=.63), suggesting a practical significance, especially for students of advanced mathematics courses in grammar schools. On average, they solved more tasks with the help of the AI, thereby also indicating a broader knowledge transfer and an increase in learning performance. However, it is important to note that both the test group of the basic courses as well as all the control groups, which used conventional keyword searches without AI, showed only minimal improvements in scores without any significance of learning improvement. This implies that the AI learning assistant indeed enhances task completion efficiency compared to traditional methods.

Deriving from these, Hypothesis 2 is also partly supported, considering the average amount of videos watched and the frequency of their usage for assistance. Among grammar school students, regardless how many videos they watched (either 1-3 or >4) those who used the AI learning assistant showed a significant improvement in their learning performance at a medium effect size. Contrastingly, control groups who did not have access to the learning assistant, but watched 1-3 instructional videos, still achieved higher total scores, whereas students who watched >4 videos worsened their results. Secondary school students showed the same pattern in support of the use of AIEDN.

Across all groups, both a high and a low utilization of endorsing video content improved their total scores, showing an increased significance for a higher number of videos watched. Interestingly, contrary to the other school type, secondary school students also benefited in the same way using the simulation prototype, although one opposite scenario could be observed: the more the number of videos used decreases while using the simulation, the more the average score worsens. This suggests that the AI assistant may help students build deeper knowledge, retain and apply it effectively for both, grammar schools and secondary schools with an increased number of educational videos watched. However, for the most part, no significant improvement could be accomplished without the underpinning of AI, regardless of the school type.

Therefore, Hypothesis 3 also received limited support in the study. Grammar school students who self-assessed a previously higher frequency of using learning assistants like AIEDN, demonstrated a significant improvement in their learning outcomes through it, suggesting that the AI tool helped with knowledge retention. In contrast, having no prior experience with these tools, did not cause any change in performance. Strikingly, secondary school students revealed the reverse scheme, having no learning enhancement through the preliminary use of learning assistants as well as AIEDN, but instead achieving lower overall scores. On the other hand, secondary school students with no prior exposure, improved their results significantly positively.

In summary, grammar school students who frequently watched instructional videos benefited more than average from the use of the AIEDN AI assistant and were able to significantly improve their scores. On the other hand, students from secondary schools who rarely or never used video tutorials before using AIEDN, improved significantly. In the overall evaluation, however, all students who were able to use the AI assistant improved.

Students who frequently studied with instructional, respectively learning videos, as well as those who rarely studied with them, both benefited from using the tool.

To conclude, the AIEDN AI learning assistant facilitates deeper knowledge acquisition for high(er)-performing students. Its effectiveness varies across performance levels. The study provides valuable insights into the potential benefits and limitations of AI-based learning support systems. Future research could delve deeper into the specific learning strategies and behaviors of students using AI assistants and explore ways to optimize AI-driven personalized learning experiences for a broader range of learners.

Limitations

In addition to the (significant) increase in learning performance and motivation, the AI assistant also showed limitations. According to the pre-survey, the most common place for the students surveyed to study was at home, due to the low noise level. This means that the study carried out in the school context at home was subject to extraneous influences that affected the study of performance improvement due to the increased stress factor of external observation, the time constraints and the number of tasks.

Furthermore, it could be observed that especially parts of the groups of students from the participating secondary schools did not actively work on the assigned tasks for a large part of the survey period. This can be partially explained through the teachers' feedback that the videos conceptualized and recorded by Daniel Jung are primarily aimed to support grammar school students. Therefore, they use mathematical terminology that does not accurately reflect the teaching concepts and materials used within secondary schools. Other plausible variables that could have influenced the study's results in this context are the students' daily routines, their different learning histories, their affinity for different kinds of technology and their general motivation to participate due to voluntariness, respectively a lack of incentives (e.g. in the form of a real exam grade).

On account of the complexity of the tasks, the study did not examine how much faster the students solved the tasks in the second session compared to the first. Thus, it is not possible to say for certain whether there was an improvement in individual subject areas. It is also unclear how much time was spent on each task and whether transfer knowledge was built between them, as they could be completed flexibly rather than sequentially. Due to the lack of an implementation of a URL blocker or similar tools, it was not possible to fully control whether respondents were able to leave the prototype on the device during the test and complete the task elsewhere.

The limitations of the prototype available at the time of the study, due to its lean design, meant that adaptive learning was supported in a limited way: visually and aurally in the form of videos played by the learning assistant, but not suitable for text-only learners. This circumstance may also have led to stimulus overload, as too many videos were offered in a given time frame, with no clear limit to the number of tasks that could be completed. Another limitation was the need for a stable internet connection, as the prototype is browser-based. Unfortunately, the provision of a stable internet connection could not be guaranteed in all schools, as there are always bottlenecks in network coverage nationwide (Bundesnetzagentur, 2022; Statista, 2023).

Directions for Future Research

Based on the research results, it can be stated that due to the high level of abstraction of the tasks, respectively the video content played, no significant correlation could be found that would indicate an improvement in the learning of secondary school students. Further research is needed on the adaptation of specific content and the selection of content creators, especially regarding age and performance level. A long-term study, e.g. the time period of one school year, would be a suitable framework. This would have several advantages: Individual learning gains and perspective learning gains could be achieved through the ability to more specifically capture the learning environments and associated confounding factors and control for their effect sizes, such as the on-site external observation that took place in this study versus distracting factors in the home environment. This opens up the possibility of testing the effectiveness of the AI assistant using real-world assessment scales and criteria, such as graded homework vs. exams. This also allows for a more sustainable adaptation to the platform, as the goal is gradual learning and thus continuous topic-specific improvement.

AI has established itself as an indispensable tool in higher education. In this context, Popenici & Kerr (2017) and Schön et al. (2023) unanimously point to the many possibilities that AI solutions open up for teaching and learning. These new opportunities are particularly valuable at a time when traditional forms of teaching are affected by declining student enrollments (Seaman et al., 2018) and there is a need to provide educational opportunities in a cost-effective manner (Kim et al., 2020).

The use of AI-enhanced teaching methods opens up a variety of opportunities that benefit both instructors and students. Rudolph et al. (2023) highlight how AI assistants can reduce the burden on instructors by automating assessment, management, and feedback processes. Students benefit from individualized learning paths and personalized learning instructions that can be developed by AI assistants based on existing learning materials, prior knowledge, skills, and pace (Schön et al., 2023).

While AI solutions are already being successfully deployed in higher education, the real need and potential for AI is most evident in primary and secondary education (K-12). It should be noted, however, that integrating AI into K-12 learning contexts presents unique challenges that require careful design considerations. As Zhou et al. (2020) point out, the needs and challenges in K-12 schools are different from those in higher education environments. Aspects such as engagement and scaffolding play a key role. Adapting AI learning tools and curricula to the K-12 context is therefore critical to meeting specific needs and maximizing the benefits of AI in schooling.

One promising approach to engage secondary students in AI while supporting their learning is the integration of gamification. The use of game-like elements can help increase student motivation and engagement while helping them learn important AI concepts (Zhou et al., 2020). Although AI in education (K-12) has already made great strides, the literature to date has mainly focused on providing students with a general understanding of AI and introducing them to the use of this technology. Concrete use cases, especially in education, are still largely unexplored. Therefore, there is a clear need to focus more on developing curricula and tools that not only teach students AI skills, but also demonstrate concrete use cases in a school context.

Acknowledgements

We would like to thank all who have contributed to the development of the AIEDN study. First, we would like to thank our cooperation partner - thingsTHINKING GmbH in cooperation with Daniel Jung Media GmbH - for the successful cooperation in our joint research project and the constructive feedback during the development of the AIEDN AI Learning assistant. We would also like to thank the Ministry of Economics, Labour, and Tourism of Baden-Württemberg for their trust and financial support. Furthermore, we would also like to thank our colleagues at the Institute for Applied Research (IAF) and the Institute for Applied Artificial Intelligence (IAAI) at Stuttgart Media University for their advice and support. Special thanks go to all headmasters, teachers, and participating students of the secondary schools involved in the study, as well as all experts who were available for interviews. We would also like to thank Selina Donat, Carina Simone Weber, Sven Kottmann, Jana Cuntz, Julia Schallmeir, Dominik Maier, Marie-Sophie Michels, Marvin Clauß, Laura Wohnus, Sarah Metzger, and Philipp Dreyer for their active support in carrying out the research and for publicizing our research project.

References

Ahmad, S. F., Alam, M. M., Rahmat, Mohd. K., Mubarik, M. S., & Hyder, S. I. (2022). Academic and Administrative Role of Artificial Intelligence in Education. *Sustainability*, *14*(3), 1101. https://doi.org/10.3390/su14031101

Alam, A., & Mohanty, A. (2022). Business Models, Business Strategies, and Innovations in EdTech Companies: Integration of Learning Analytics and Artificial Intelligence in Higher Education. *2022 IEEE 6th Conference on Information and Communication Technology (CICT)*, 1–6. https://doi.org/10.1109/CICT56698.2022.9997887

Bundesministerium für Bildung und Forschung (2024). Zusammenarbeit von Bund und Ländern, (2015). Retrieved from: https://www.bmbf.de/bmbf/de/bildung/zusammenarbeit-von-bund-und-laendern/zusammenarbeit-von-bund-und-laendern_node.html

Bundesnetzagentur. (2022). *Jahresbericht Telekommunikation 2022*.

Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial Intelligence trends in education: A narrative overview. *Procedia Computer Science*, *136*, 16–24. https://doi.org/10.1016/j.procs.2018.08.233

Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in Education: A Review. *IEEE Access*, *8*, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

Ciolacu, M., Tehrani, A. F., Binder, L., & Svasta, P. M. (2018). Education 4.0 - Artificial Intelligence Assisted Higher Education: Early recognition System with Machine Learning to support Students' Success. *2018 IEEE 24th International Symposium for Design and Technology in Electronic Packaging (SIITME)*, 23–30. https://doi.org/10.1109/SIITME.2018.8599203

Engels & Schüler. (2020). *Bildung digital? Wie Jugendliche lernen und Schulen lehren*. IW-Trends – Vierteljahresschrift zur empirischen Wirtschaftsforschung, ISSN 1864-810X, Institut der deutschen Wirtschaft (IW), Köln, Vol. 47, Iss. 2, pp. 89-106, https://doi.org/10.2373/1864-810X.20-02-05.

Findeisen, S., Horn, S., & Seifried, J. (2019). Lernen durch Videos – Empirische Befunde zur Gestaltung von

Erklärvideos. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, 16–36. https://doi.org/10.21240/mpaed/00/2019.10.01.X

Hartley, C. A., Nussenbaum, K., & Cohen, A. O. (2021). Interactive Development of Adaptive Learning and Memory. *Annual Review of Developmental Psychology*, *3*(1), 59–85. https://doi.org/10.1146/annurev-devpsych-050620-030227

Kim, J., Merrill, K., Xu, K., & Sellnow, D. D. (2020). My Teacher Is a Machine: Understanding Students' Perceptions of AI Teaching Assistants in Online Education. *International Journal of Human–Computer Interaction*, *36*(20), 1902–1911. https://doi.org/10.1080/10447318.2020.1801227

Kokoç, M., IIgaz, H., & Altun, A. (2020). Effects of sustained attention and video lecture types on learning performances. *Educational Technology Research and Development*, *68*(6), 3015–3039. https://doi.org/10.1007/s11423-020-09829-7

Lämmel, U., & Cleve, J. (2020). *Künstliche Intelligenz*. Carl Hanser Verlag GmbH & Co. KG. https://doi.org/10.3139/9783446463639

Lee, I., & Perret, B. (2022). Preparing High School Teachers to Integrate AI Methods into STEM Classrooms. *Proceedings of the AAAI Conference on Artificial Intelligence*, *36*(11), 12783–12791. https://doi.org/10.1609/aaai.v36i11.21557

Liu, Y., Saleh, S., & Huang, J. (2021). Artificial Intelligence in Promoting Teaching and Learning Transformation in Schools. *International Journal of Innovation, Creativity and Change*, 891–902. https://doi.org/10.53333/IJICC2013/15369

Lloyd, S. A., & Robertson, C. L. (2012). Screencast Tutorials Enhance Student Learning of Statistics. *Teaching of Psychology*, *39*(1), 67–71. https://doi.org/10.1177/0098628311430640

Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., Hinz, O., Morana, S., & Söllner, M. (2019). AI-Based Digital Assistants: Opportunities, Threats, and Research Perspectives. *Business & Information Systems Engineering*, *61*(4), 535–544. https://doi.org/10.1007/s12599-019-00600-8

Ouyang, F., Zheng, L., & Jiao, P. (2022). Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. *Education and Information Technologies*, *27*(6), 7893–7925. https://doi.org/10.1007/s10639-022-10925-9

Palacios Hidalgo, F. J., Huertas Abril, C. A., & Gómez Parra, M. ª E. (2020). MOOCs: Origins, Concept and Didactic Applications: A Systematic Review of the Literature (2012–2019). *Technology, Knowledge and Learning*, *25*(4), 853–879. https://doi.org/10.1007/s10758-019-09433-6

Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, *12*(1), 22. https://doi.org/10.1186/s41039-017-0062-8

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching*, *6*(1). https://doi.org/10.37074/jalt.2023.6.1.9

Schleiss, J., & Göllner, S. (2022). *Rahmenbedingungen für Künstliche Intelligenz in Educational Technology*. https://doi.org/10.18420/DELFI2022-WS-40

Schön, E.-M., Neumann, M., Hofmann-Stölting, C., Baeza-Yates, R., & Rauschenberger, M. (2023). How are AI assistants changing higher education? *Frontiers in Computer Science*, *5*, 1208550. https://doi.org/10.3389/fcomp.2023.1208550

Seaman, J. E., Allen, I. E., & Seaman, J. (2018). *Grade Increase: Tracking Distance Education in the United States*. Babson Survey Research Group.

Staatsministerium für Kultus. (2016). *Das Abitur am allgemeinbildenden Gymnasium*.

Statista. (2023). *Fläche mit Mobilfunkversorgung nach Technologie 2023*. Retrieved from:https://de.statista.com/statistik/daten/studie/1346916/umfrage/flaeche-in-deutschland-mit-mobilfunkversorgung-nach-technologie/

Van Der Meij, H., & Van Der Meij, J. (2014). A comparison of paper-based and video tutorials for software learning. *Computers & Education*, *78*, 150–159. https://doi.org/10.1016/j.compedu.2014.06.003

Vonderau, P. (2016). The video bubble: Multichannel networks and the transformation of YouTube. *Convergence: The International Journal of Research into New Media Technologies*, *22*(4), 361–375. https://doi.org/10.1177/1354856516641882

Zhou, X., Van Brummelen, J., & Lin, P. (2020). *Designing AI Learning Experiences for K-12: Emerging Works, Future Opportunities and a Design Framework*. https://doi.org/10.48550/ARXIV.2009.10228

Author Information

Vivian Kretzschmar
https://orcid.org/0009-0008-7692-4847
Stuttgart Media University
Nobelstrasse 10; 70569 Stuttgart
Germany
Contact e-mail: *kretzschmarvivian@gmail.com*

Alica Sailer
https://orcid.org/0009-0003-0217-9630
Stuttgart Media University
Nobelstrasse 10; 70569 Stuttgart
Germany

Michael Wertenauer
https://orcid.org/0009-0003-1212-5273
Stuttgart Media University
Nobelstrasse 10; 70569 Stuttgart
Germany

Prof. Dr. Jürgen Seitz
https://orcid.org/0009-0001-0465-1447
Stuttgart Media University
Nobelstrasse 10; 70569 Stuttgart
Germany