Using AI Chatbot for Math Tutoring

Antonín Jančařík

Department of Mathematics and Mathematical Education
Faculty of Education, Charles University
M. D. Rettigové 4, 116 39, Prague 1, Czech Republic
E-mail address: antonin.jancarik@pedf.cuni.cz
ORCID: https://orcid.org/0000-0003-3331-2396

Jakub Michal

Department of Mathematics and Mathematical Education
Faculty of Education, Charles University
M. D. Rettigové 4, 116 39, Prague 1, Czech Republic
E-mail address: cjakub@email.cz
ORCID: https://orcid.org/0000-0003-3522-8411

Jarmila Novotná

Department of Mathematics and Mathematical Education
Faculty of Education, Charles University
M. D. Rettigové 4, 116 39, Prague 1, Czech Republic
&
CeDS Univerité Bordeaux
351 cours de la Libération, 33405 Talence CEDEX, France
E-mail address: jarmila.novotna@pedf.cuni.cz
ORCID: https://orcid.org/ 0000-0002-5306-2315

Abstract

Aim. The aim of this paper is to present the results of a pilot validation of an artificial intelligence chatbot for mathematics tutoring. The subject of the analysis is the behaviour of the students in the course, in particular their approach to solving problems without external motivation through scoring or other assessment.

Methods. The analysis of student behaviour was carried out on the basis of records in the application database. The subject of the analysis was not only the pupils' responses, but also the time taken to produce them and the overall trajectory through the course.

Results. The results of the analysis show that students who choose to join the course, work intensively in the lessons. Most pupils solve more than one basic task. A significant proportion of pupils also use the prepared help and instructional videos. The analysis of the solution time allowed identification of a group of pupils who do not even read the problems and only confirm the results.

Conclusion. The pilot validation confirmed that the chat-bot format is close to the learners, and they are able to use the courses created in this environment. Solution time analysis suggests that this parameter can be used to analyse learner behaviour and system control. Further analyses in this area are needed.

Keywords: artificial intelligence, tutoring, mathematics education, time analysis, chatbots

Introduction

Recently, there has been significant progress in the field of artificial intelligence (AI) across various aspects of human life. The rapid advancement of AI systems has sparked widespread discussions in society, including its impact on education (for example, Rahman, 2021).

This paper presents the findings of a research study conducted as part of the AI Assistant for Pupils and Teachers project. The project aimed to develop an AI chatbot system specifically designed to assist students in learning mathematics (Jančařík, Novotná et al., 2022; Jančařík , Michal et al., 2022). The application involves students interacting with an AI chatbot, which presents them with pre-prepared mathematical problems to solve. It is important to note that the AI's role is not to generate or prepare mathematical content, but rather to facilitate communication with the student beyond the assigned tasks. In this paper, we will present an initial analysis of the utilisation of mathematical content and propose some system modifications based on the collected data.

Purpose and Objective

The term tutoring in this paper refers to private supplementary tutoring in school mathematics. Despite the fact that there are a number of different forms of tutoring, the prevailing form is the form of private supplementary tutoring, i.e., paid tutoring focusing on content from school lessons or on preparation for entrance exams (Novotná, 2015). There are many articles dedicated to study of tutoring and online education in different settings (i.e. Al-Hawamdeh 2022, Kildė, 2023, Pituła, 2022 or Weiss et al., 2023). Private supplementary tutoring is not available to all students to the same extent, differences are e.g. in family socio-economic. Tutored pupils usually achieve better results (Safarzyńska, 2013). The fact that there is direct contact between the tutor and the pupil offers a more balanced environment for all pupils interested in private tutoring and can decrease the above-mentioned inequalities (Jančařík et al., 2023). In this way, AI can contribute to an increase in the frequency of use of online tutoring – the tutor can be replaced in the selection of the study trajectory, in evaluation of results, as well as in communication with the student (Alhossaini & Aloqeely, 2021, Jančařík et al., 2023).

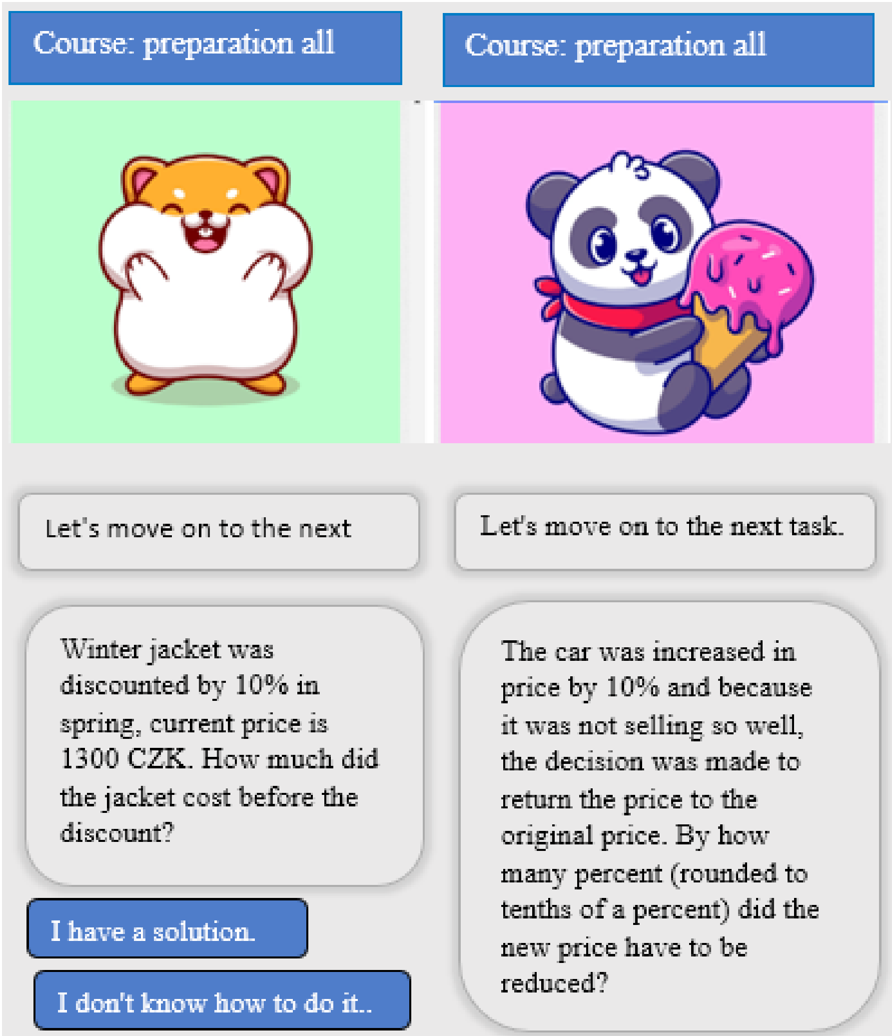

Figure 1
Example of tutoring app communication

Source. Own research.

The goal of this article is to present issues related to the framework of online mathematics tutoring with AI in the system of independent individual teaching/learning. We limit our considerations to the domain of school mathematics. The presented research project aims to create, experimentally practice and analyse pupils' use of a system proposed especially to individual tutoring when solving problems. The following research questions are addressed: Does the proposed course format meet the needs of the students? How the chosen procedures should be modified to better serve this purpose?

The tutoring program developed for this study consists of a series of thematic units, specifically designed to cover the topics necessary for successfully passing the entrance state mathematics exams (Jančařík, Michal et al., 2022b). Each topic has undergone thorough examination and has been divided into multiple lessons. These lessons are structured to include problems at three difficulty levels (easy, basic and difficult), along with an instructional video and supplementary materials. It is important to note that the progression of lessons is not linear but is represented by the following flow chart.

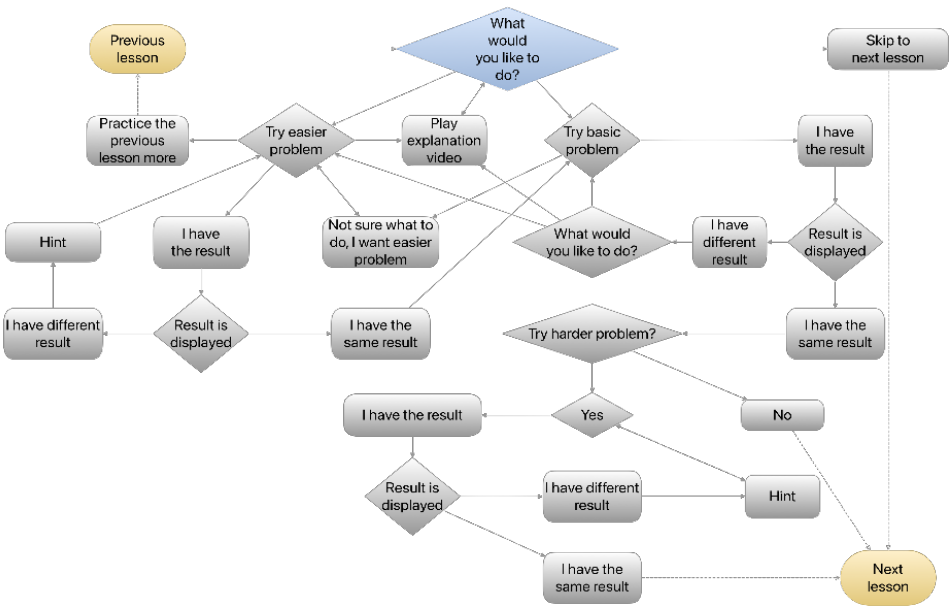

Figure 2
Scheme of tutoring app workings
Source. Own research.

The learner has the freedom to choose their own progression through the course. By integrating the prepared materials with the AI chatbot, learners have the opportunity to ask questions or engage in communication on any topic at any point during the course. However, it should be noted that the app's primary function is not to facilitate free-form conversations, thus the communication possibilities are limited. Nevertheless, through an integrated algorithm, the application is capable of recognising specific queries and providing relevant answers, along with links to the corresponding sources of information.

A unique scoring system is implemented within the application, which presents only the correct solution for review once the learner has solved a problem (Jančařík, Novotná et al., 2022a). This approach aims to distinguish tutoring from regular classroom teaching and alleviate the pressure associated with scoring, reducing the likelihood of random guessing of results. Consequently, one of the primary goals of the experimental validation is to confirm the theoretical assumptions and assess whether the course records can be utilised to differentiate pupils who actively solve problems from those who simply confirm correct solutions. In this analysis, we will examine the time traces of course progression to achieve these objectives.

Research Methodology

The analysis of the course passage was carried out on the basis of the course passage records in the application database. The second lesson dedicated to the topic of percentages was selected for analysis. This is the second, advanced lesson on this topic. The lesson contains the following three problems:

Easier problem: The book was discounted from 1800 CZK to 1044 CZK before Christmas. What was the discount in percentage?

If a student asks for a hint, the following text is sent by the system:

If the book cost 1800 (100 %) at first, what percentage is 1044? How many percent less?

Basic problem: The car was made 10 % more expensive. Because it was not selling so well, the decision was made to return the price to the original. By how many percent (rounded to tenths of a percent) did the new price have to decrease?

If a student asks for a hint, the following text is sent by the system:

Think about what was going on with the price. In the beginning, the car cost x crowns. After the increase, it cost 1.1x crowns. What do you need to multiply it by to make it cost x again? Then you can easily calculate the percentage by which it was cheapened.

Harder problem: A smaller block was cut from the block, the edges of which were one-fifth the original length. What percentage of the original block was removed?

If a student asks for a hint, he or she will receive the following text in the system: Mark the original volume of the cube V = abc. Determine the volume of the smaller block. Then you just need to determine what percentage of the large one it is.

In addition to the assignments, hints and answer checks, the lesson also includes a tutorial video showing a sample procedure for solving problems with the topic of percentages.

Data analysis was carried out in two phases. In the first phase, a basic descriptive analysis of students' behaviour in the course was conducted. The second phase then looked in more detail at how pupils answered each question, particularly in relation to how much time they spent calculating and checking their answers.

Table 1 shows an example of a recording of one student's progress through the lesson. In addition to the question and answer, the time elapsed between the display of the question and the student's answer is calculated for each step. The student in question started the lesson by viewing the tutorial videos. He then proceeded to solve a simple problem, which he solved in 18.5 seconds, but then had a problem solving the spreading problem despite the help display, and repeatedly returned to the basic problem. He eventually did not proceed to solve the difficult problem and chose to proceed to the next lesson.

Table 1

Examples of progressions through lesson

ID	*Chatbot*	*Time*	*Student*
1602	Decide what to do next	17,01	Video with explanation
1602	Decide what to do next	133,56	I want to try a simple task.
1602	The book was discounted from 1800 CZK to 1044 CZK before Christmas. What was the discount in percentages?	18,52	I have the result.
1602	Did you get the same result as me?	1,62	42% discount
1602	The car was made 10% more expensive. Because it was not selling so well, the decision was made to return the price to the original. By how many percent (rounded to tenths of a percent) did the new price have to fall?	7,87	I do not know how to proceed.
1602	I guess something was wrong with your solution. What do you want to do now?	3,94	I want to try a simple task.
1602	The book was discounted from 1800 CZK to 1044 CZK before Christmas. What was the discount in percentages?	1,27	I have the result.
1602	Did you get the same result as me?	0,23	42% discount
1602	The car was made 10% more expensive. Because it was not selling so well, the decision was made to return the price to the original. By how many percent (rounded to tenths of a percent) did the new price have to fall?	4,75	I have the result.
1602	Did you get the same result as me?	0,81	I have a different result.
1602	I guess something was wrong with your solution. What do you want to do now?	0,93	I want to try a simple task.
1602	The book was discounted from 1800 CZK to 1044 CZK before Christmas. What was the discount in percentages?	0,58	I have the result.
1602	Did you get the same result as me?	0,23	42% discount
1602	The car was made 10% more expensive. Because it was not selling so well, the decision was made to return the price to the original. By how many percent (rounded to tenths of a percent) did the new price have to fall?	1,16	I have the result.
1602	Did you get the same result as me?	0,81	About 9,1 %
1602	You've completed the basic task; do you want to solve more complex one?	1,04	No
1602	Are we moving on?	0,69	Sure

Source. Own research.

Basic Statistics

121 pupils entered the course. Of these, 66 (57%) solved at least one problem.24 (19.8%) students decided to skip the lesson on percentages immediately. 14 (11.6%) pupils first watched the video with explanation or started to interact in the chat and only then left the lesson. 12 (9.9%) students left the lesson without responding. 1 student reset the entire lesson. 4 students chose to switch to an assignment but then communication was interrupted.

In the following, we focus on the analysis of the solutions of the 66 students who solved at least one problem in the course.

23 pupils solved all three problems, 13 pupils solved the easier and basic problem, 20 pupils solved the basic and more difficult problem. 2 pupils finished the work on the easy problem, 8 pupils solved only the basic problem (see Figure 3).

From the results it can be seen that when pupils proceeded to solve the problems in the text, most of them chose to solve multiple problems.

Figure 3

Number of pupils solving assigned tasks in the lesson on percentages

Source. Own research.

Easier Problem

The simple problem was solved by 38 students.25 of them chose to solve the simple problem as their first choice, 4 students first watched the video and then moved on to the simple problem.

Of the 38 students who solved the problem, 26 declared after submitting the problem that they had solved the problem, of which 18 confirmed that they had the correct solution. Of the remaining 8, 4 solved the problem after taking the hint, 2 solved the problem after watching the video and 2 did not solve the problem. Of the remaining 12, 7 students asked for help and 6 had already solved the problem afterwards and 1 had not solved the problem. 2 students solved the problem after watching the video and 3 first interacted with the chatbot, 2 solved the problem and 1 did not. In total, 34 of the 38 students who solved the problem eventually declared the correct solution. As the following analysis will show, this result is biased by the fact that some students only declared that they had a solution without solving the problem (cf. Vonkova at al., 2022).

Some pupils (7) asked for help or declared that they did not have the correct solution.

Basic Problem

The basic problem was solved by 63 pupils. 42 of them started with this task. 41 pupils declared that they had solved the problem after it was displayed, 27 subsequently confirmed their agreement with the displayed result. Two pupils did not complete the problem. The remaining 12 pupils solved the problem either after viewing the clue (9) or after viewing the clue and the video (3). A total of 6 pupils viewed the video for this problem.

Harder Problem

A total of 40 pupils solved the difficult problem. After the display, 32 answered that they had solved the problem and 22 of them confirmed that they had solved it correctly. After the clue, all pupils indicated that they had solved the problem, with one pupil displaying the clue twice.

Data Analysis Based on Solution Time

On the one hand, basic data analysis shows that a large proportion of students self-report if they are unable to solve a problem and ask for help, either in the form of a hint or a tutorial video. On the other hand, the data obtained indicate that some of the pupils do not solve the problems but only mark the correct answers. An indicator that points to such behaviour is the time elapsed between the display of the question and the answer. If this time is in the tens of seconds, it is clear that these are times that are completely inadequate not only for solving the problem but also for simply reading it. We therefore performed an analysis of the solution times.

The aim of this analysis is to obtain data for further adjustments to the system, especially adjusting it to react adequately to the situation when the student stops solving the problems and starts only confirming the correct answers. For the easy problem, pupils who asked for help before seeing the answer (saying they did not know how to solve) or the video gave up solving independently after 17.6 seconds on average, for the basic problem after 8 seconds and for the difficult problem after 30.4 seconds.

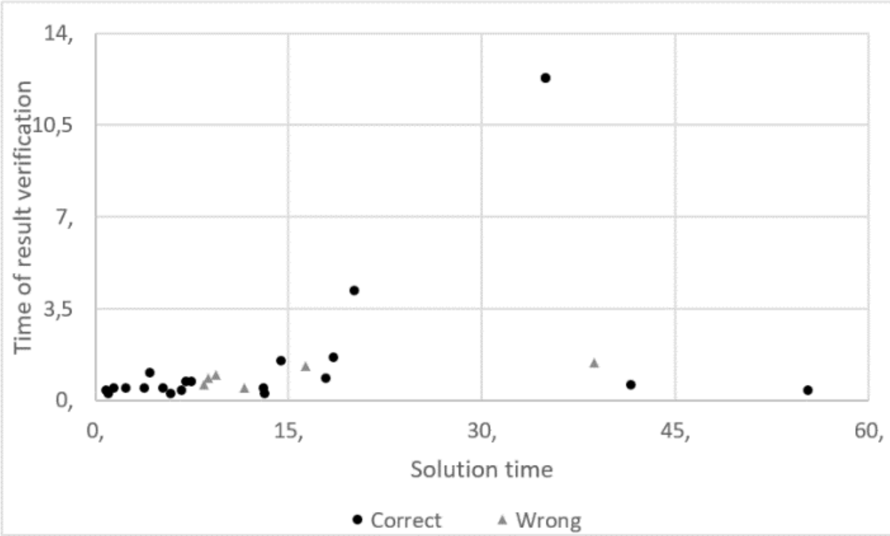

Figure 4
Easy task time scores
Source. Own research.

Another characteristic that can be used to analyse the relevance of students' answers is the comparison between the time taken to solve the problem and the time taken to verify the correct answer. The following graphs (Figures 4, 5, and 6) compare these times for the basic, intermediate, and difficult problems. The graphs show that some students probably did not solve the problem at all and only confirmed the correctness of the answer (high density of points in the lower left part of the graphs). At the same time, we also see that the time needed to confirm the correctness of the answer is usually less than 2 seconds. It was observed (for the repeated problem solving) that the time needed to confirm the answer was much longer and corresponded more to the time needed to solve the problem. In this case, we assume that the student solved the problem incorrectly and then recalculated the problem before confirming the correct answer. In further modifications we will try to set up an adequate response to this situation as well, e.g., by asking whether this was indeed the case.

We also note in Figure 5 that despite the expectation that students who answer quickly on the baseline problem will consider 10% as the correct answer and thus have a wrong answer, there are only 4 acknowledged wrong answers within 25 seconds. Since the problem is not easy to solve at first sight, this result also indicates that these students probably did not solve the problem at all.

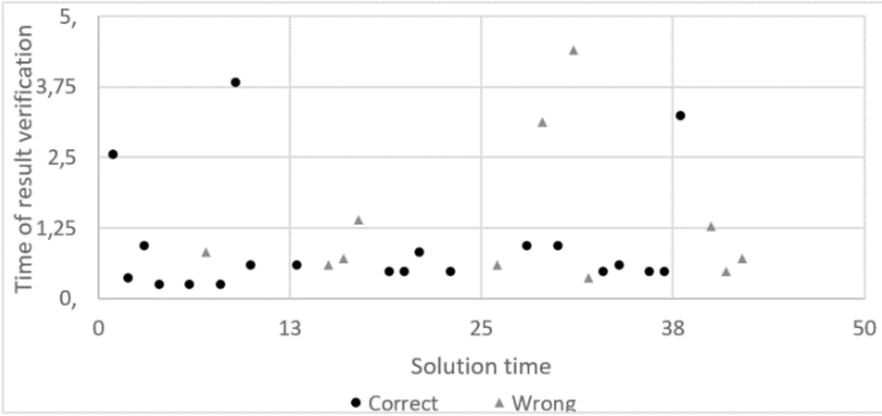

Figure 5
Basic task time scores

Source. Own research.

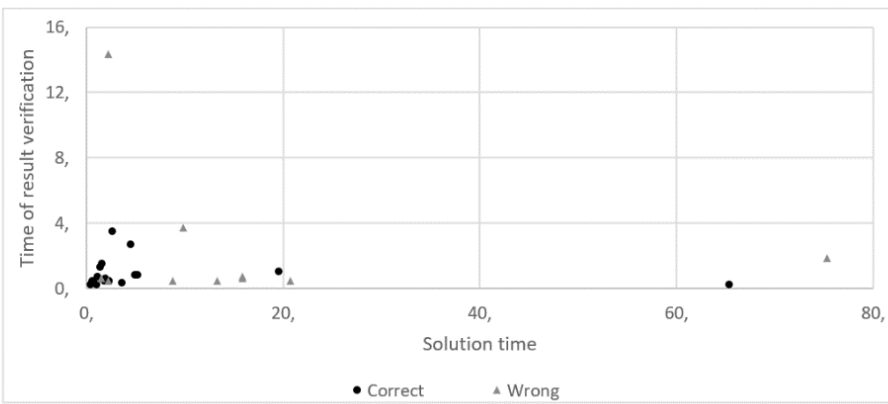

Figure 6
Challenging task time scores

Source. Own research.

When comparing the solution times for the easy, basic and difficult problems, it was found that students generally spent the shortest time on the difficult problem (up to 21 seconds with two exceptions) and also confirmed quickly, i.e., did not try to understand the problem with the solution shown. The fact that pupils solved the problem relatively quickly can be explained either by the possibility that pupils gave up on the solution earlier, because they knew the problem was difficult and did not believe they could do it, or as a consequence of the fact that the difficult problem was already mostly accessed by pupils who did not read the options and just clicked randomly. Conversely, pupils who solved honestly often skipped the difficult task. Similarly, the short time devoted to the displayed solution can be explained.

Conclusion

The popularity of tutoring has been on the rise among students (Novotná, 2015). Tutoring is also impacting the education system itself. Students tend to value a more personalized approach to their tutoring, aiming for greater individualization of their learning experience and a broader scope of knowledge. Individualisation of approach is an area of focus in our research on the use of AI in education.

We consider the use of artificial intelligence to be important because it makes tutoring free to anyone, anytime. The system that is being developed and that we present in this paper works with pre-prepared mathematical content and an assessment system that puts the responsibility for checking the results on the learner. The pilot research on pupil behaviour in the course mainly yielded the following findings:

- Students are able to use AI Chabot as a tutoring tool when they already access the prepared course content, solving usually more than one problem.
- The prepared system of hints and tutorial videos is used by students as a support for solving tasks and helps them to master the discussed material.
- The single correct answer system is proven to work, as a large number of pupils ask for help and hints instead of simply confirming the correct answer.
- Some pupils simply confirm the correct solutions during the solution process, without solving the problems or, in some cases, even reading them. The implemented analysis of the answer times is able to detect such solutions, allowing the addition of an adequate response to the system.
- Contrary to expectations, some students repeatedly return to the easier problem. To increase the effectiveness of the course, one can consider creating more simple problems for each lesson and varying them when they are shown again.

Acknowledgements

This research was funded by the Technology Agency of the Czech Republic, grant number TAČR N. TL05000236 – AI assistant for pupils and teachers.

References

[1] Al-Hawamdeh, B. O. S. (2022). Pedagogical Problems During the Pandemic: A Study of Issues and Challenges in Teaching. *Journal of Education Culture and Society*, *13*(2), 205-220. http://doi.org/10.15503/jecs2022.2.205.220

[2] Alhossaini, M., & Aloqeely, M. (2021). Causal Analysis of On-line Math Tutoring Impact on Low-income High School Students Using Bayesian Logistic and Beta Regressions. In *IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 1-10). IEEE.

[3] Jančařík, A., Novotná, J., & Michal, J. (2022). Artificial Intelligence Assistant for Mathematics Education. In *Proceedings of the 21st European Conference on e-Learning - ECEL 2022* (pp. 143-148). Reading: Academic Conferences Ltd.

[4] Jančařík, A., Michal, J., & Novotná, J. (2022). Criteria for Classification of Digital Educational Materials and AI. In J. Fejfar & M. Flégl (Eds.), *Proceedings of the 19th International Conference Efficiency and Responsibility in Education 2022* (pp. 45-51). Czech University of Life Sciences.

[5] Jančařík, A., Michal, J., & Novotná, J. (2023) Feedback in Online Mathematics Tutoring. In J. Jovanovic, I.-A. Chounta, J. Uhomoibhi, & B. McLaren (Eds.), *Proceedings of the 15th International Conference on Computer Supported Education* (pp. 374-381). SCITEPRESS. http://doi.org/10.5220/0000173200003470

[6] Kildė, L. (2023). ESL Teachers' Approaches towards the Acceptance of Educational Technology Integration in Non-Formal Education: A Case from Kenya. *Journal of Education Culture and Society*, *14*(1), 634-649. http://doi.org/10.15503/jecs2023.1.634.649

[7] Novotná, G. (2019). Pupils' Perception of Their Understanding in Mathematics and its Connection to Private Supplementary Tutoring. In *Proceedings of CERME 11* (pp. 1477-1484). ERME. https://www.researchgate.net/publication/348579569

[8] Pituła B. (2022) Teacher in a New Reality. Will and Avatar Replace a Teacher, In B. Pituła, H. Badura, M. Mares, A. Kiepas, & J. Wolny (Eds.), *End of worlds - world of ends* (pp. 115-134). Wydawnictwo A. Marszałek.

[9] Rahman, M. M. (2021). Should I Be Scared of Artificial Intelligence? *Academia Letters,* 2536. http://doi.org/10.20935/AL2536

[10] Safarzyńska, K. (2013). Socio-economic Determinants of Demand for Private Tutoring. *European Sociological Review*, *29*(2), 139-154.

[11] Vonkova, H., Papajoanu, O., Moore, A., & Bosko, M. (2022). Questionnaire Item Response Times and Claiming Impossible Knowledge: An Analysis Using the Overclaiming Technique. In *Proceedings of IAC in Vienna 2022* (pp. 140-144). CIAE.

[12] Weiss, E., Akimjaková, B., Paľa, G., & Biryukova, Y. N. (2023). Methods of Re-Education of Specific Learning Disorders. *Journal of Education Culture and Society*, *14*(1), 185–197. http://doi.org/10.15503/jecs2023.1.185.197