The Impact of Content Design on the Effectiveness of an Educational Application

Tomáš Jeřábek, Jakub Michal and David Janda
Charles University, Faculty of Education, Prague, Czech Republic

tomas.jerabek@pedf.cuni.cz
jakub.michal@pedf.cuni.cz
david.janda@pedf.cuni.cz

Abstract: Interactive applications play a significant role in digital educational resources. Their integration has a notable impact, especially in non-formal and informal education. The success of applications depends not only on appropriate content but also on factors that increase users' motivation to use them. These aspects include gamification, activation, personalisation of content and environment, immediate feedback, multimediality or flexibility of application accessibility. The paper focuses on the optimal design of lesson content within an interactive learning application for ISCED 2 students in Mathematics. The paper aims to find relevant aspects and parameters of the application in this context and to identify their appropriate settings for more efficient learning. The study uses quantitative analysis of user behaviour within the application environment, drawing on pedagogical principles such as microlearning, content adaptivity, immediate feedback, and gamification. Lessons are divided into explanation blocks and related practice tasks. As part of the methodology, we evaluate the relationship between the design parameters of these blocks in individual lessons and users' progress through lessons. To assess the optimal setting, we examine the dependence between the overarching mathematical topic, the total amount of content, the number and type of lesson blocks, and tasks. This analysis is based on the theoretical assumption that interactive solving of educational tasks represents the so-called active screen time, which requires the user's direct cognitive involvement. Empirical data for analysis are based on system records of each student's course progress. The key indicators of interaction are analysed, particularly the time spent in individual lessons and blocks of lessons, with their structure taken into account. The synthesis of this objective data allows us to formulate recommendations to optimise educational content that maximise both students' cognitive engagement and the overall effectiveness of the learning process. Also, it helps set an appropriate methodology for further, deeper data analysis.

Keywords: Online Tutoring System, Mathematics Tutoring, Gaming, Content Design

1. Introduction

Tutoring, by which we mean private supplementary teaching conducted in parallel with schooling, is quite common in the Czech Republic. Although exact numbers are unknown, mathematics tutoring is among the most requested services (Šťastný, 2016). This learning support is usually sought to improve grades or to prepare for the high-stakes examinations (e.g. Průcha, 2015).

A free web and mobile app (Edu-AI) for primary and lower-secondary school mathematics tutoring was developed as part of an *AI-based Adaptive Learning Platform* project. It is a loose continuation of the *AI Assistant for Pupils and Teachers* project (see Jančařík et al., 2023). The current goal is to implement generative AI systems beyond chatbot use (e.g., to provide adequate feedback, hints, or personalised tasks). As we strive to make the app motivating for students and effective for learning (which does not depend solely on the quality and quantity of content), we also incorporate design principles that support long-term user motivation and cognitive engagement. Among the most frequently discussed aspects for achieving this goal are gamification elements, activation strategies, personalisation of both content and the learning environment, immediate feedback, multimedia integration, and flexible accessibility across devices and contexts. These features contribute not only to learner satisfaction but also to increased persistence and active participation in the learning process.

In its current piloting phase, the developed tutoring application comprises several teaching courses focused on lower-secondary school mathematics, specifically for 8th- and 9th-grade students. These courses stem from the *Global Proficiency Framework* (GPF, 2020), which distinguishes proficiency levels in greater depth than the *Czech National Curriculum for primary schools* (MŠMT, 2023).

In the pilot testing of the in-development application, students first completed an introductory diagnostic test, followed by eight courses focused on primary school mathematics, released weekly. This paper presents the collected dataset and provides an initial analysis of students' behaviour in the application, focusing on lesson and block characteristics that may influence engagement, particularly the time spent on different activities and possible signs of the so-called "gaming" phenomenon. The findings are intended to inform the further development of an engaging intelligent tutoring system that maintains students' motivation, provides an appropriate level of difficulty, and better adapts to students' interactions with the content.

2. Theoretical Background

Many studies have focused on tutoring or online education in mathematics (e.g., Son, 2024). Although there is no clear consensus on the effects of tutoring, and studies have yielded mixed results (Safarzyńska, 2013), several researchers have reported varying degrees of improvement in students' outcomes (Mabhamba et al., 2024; Zhang et al., 2022). Even though some studies did not observe better results of tutored students, they still observed improvements in students' well-being (Guill et al., 2020). This opportunity, however, is not available to all students, as tutoring is usually a paid service, which might widen the socio-economic gap (Safarzyńska, 2013). Not only for this reason, but also to enrich opportunities for all students, intelligent tutoring systems offered for free could improve this situation.

Even though mobile learning apps or similar environments might motivate students, improve their learning or positively influence their attitude towards mathematics (e.g., Canilao & Gurat, 2023; Taleb et al., 2014), such apps often lack the feedback, immediate assistance and communication facilitated by a real-life tutor, who can tailor the lessons to one's individual needs (Safarzyńska, 2013). There are, however, various approaches to address these restrictions (for instance, incorporating AI language models as chatbots or using animated agents [Jančařík et al., 2023; Baker et al., 2006]). Other issues tutoring systems need to take into account include students' disengagement or lack of motivation, the way curricular materials are presented or the scaffolding methods used.

2.1 Microlearning

Microlearning builds on principles from psychology and pedagogy and is often associated with e-learning (Kapp & DeFelice, 2019). Microlearning is usually explained in terms of either the short *time* during which learning occurs or the small *size* of the targeted teaching units (e.g., Cates et al., 2017). In both instances, learning is organised into small, manageable units of time or information, which is believed to help prevent mental fatigue by reducing demands on working memory. On small mobile screens, such an approach is welcomed (sometimes called *mobile microlearning*), as displaying too much information can lead to cognitive overload (Lee, 2023). Further advantages are also reported in the literature – greater flexibility, a focus on the knowledge to be learned, longer and easier information retention, and stronger engagement (Kadhem, 2017; Kapp & DeFelice, 2019; Lee, 2023; Şahin & Kirmizigül, 2026; Taylor & Hung, 2022).

The effectiveness of microlearning was, for instance, observed in research conducted by Závodná et al. (2024). Students achieved better results with microlearning approaches and were more motivated towards their work than those in the control group. Students further appreciated the visual features, gamification, and interactive quizzes. The effectiveness of microlearning was also observed in mathematics education. Şahin and Kirmizigül (2026) reported that students found microlearning mathematics courses engaging and flexible, and that they appreciated the opportunity to review the content. They also recommend integrating microlearning with face-to-face teaching.

The so-called micro-lectures usually employ different strategies when being made. Based on a literature review, Tang et al. (2022) identify several strategies, including assignments, problem-based learning, feedback interactions, connecting learning to objects from students' daily lives, knowledge strengthening, and experimental demonstrations.

2.2 Gaming the System Issue

One difficulty with tutoring systems is students' disengagement over time (Arroyo et al., 2007; Baker et al., 2006). This can manifest as skimming through the app, immediate use of hints without trying to solve or even read anything offered, or attempts to exploit the system rather than learn – a phenomenon known as „gaming" (Arroyo et al., 2007; Baker et al., 2006). Even though there are instances when gaming might not be problematic (in tasks students know well), it is unwanted behaviour in tasks students struggle with (Baker, 2006).

To address this difficulty, different approaches might be taken – from user-friendly application design and gamification features to self-reflection or self-monitoring systems. For instance, Arroyo et al. (2007) found that non-invasive interventions (such as providing study tips or a performance chart) reduced quick-guessing behaviour and increased students' motivation and re-engagement with a geometry tutoring system. As such behaviour might stem, for instance, from frustration at being unable to solve a challenging task, it is important that extra hints or help be offered (Baker et al., 2006). On the other hand, Arroyo et al. (2007) note that some intrusive attempts to remedy such behaviour (e.g., pop-up windows with information) might be

counterproductive, increasing students' frustration. Thus, they recommend making students aware of the consequences of such actions and their effects on learning progression, without strictly blocking this approach. For instance, Baker et al. (2006) were successful with such a positive approach, using an animated agent (a dog with a range of emotions) designed to improve gaming students' learning outcomes without significantly restricting the system for students who did not participate in gaming. It was done by praising non-gaming students, by challenging students to work carefully when gaming was identified, or by giving additional exercises designed to help gaming students cover the material they had skipped. This led to fewer gaming students (due to the dog's sole presence, regardless of the emotions displayed) and to students who struggled to catch up with their peers on the content.

2.3 Application Content Design

Based on the review above, several principles were adopted in the design of the application Edu-AI. The app is intended for tutoring mathematics and covers all lower secondary school topics. Among others, the microlearning principle was implemented. To increase engagement and prevent gaming, gamification features (collecting experience points that can be used to „buy" in-app avatars or visual modifications) and a visually appealing design were applied. Flexibility is also enhanced by the option to save blocks or quiz questions that a student finds problematic for later review. To accommodate students' diverse personal learning needs (e.g., additional personalised hints, feedback, or questions), an AI chatbot will be integrated in the next step of app development.

The content is based on a thematic course structure. Each course is divided into one or more teaching sections (lessons) intended for 10-minute study sessions. Lessons consist of building units (blocks) of various types–pre-prepared explanatory text blocks, blocks with illustrations, example blocks, or task blocks (see Figure 1). The explanatory blocks focus on helping students understand the concepts they are expected to know as they leave primary school and on connecting the material to what students know from day-to-day experience. In task blocks, students have access to two-step hints and immediate feedback, usually explaining the anticipated mistake (see Figure 1). The correct solution is also displayed before continuing to the following block.

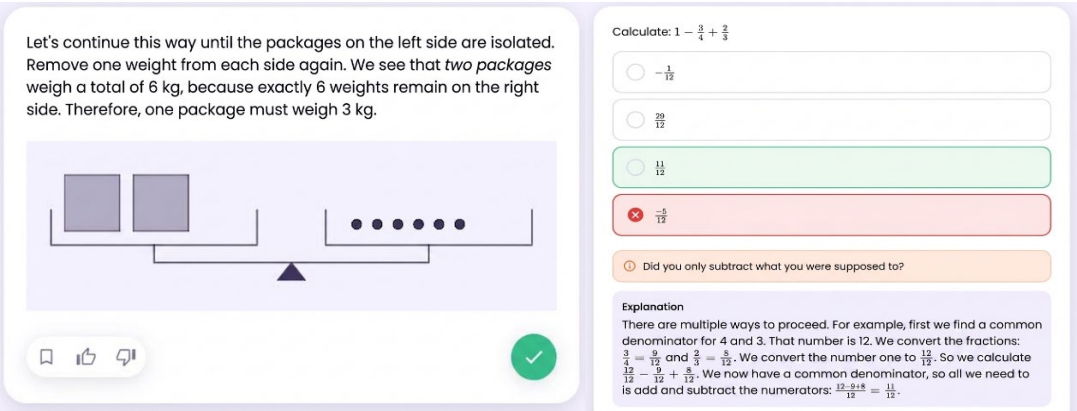


Figure 1: Example of the content block (left) and the task type block (right) (translated)

3. Materials and Methods

Testing of the application took place in March and April 2026 and was conducted in two ways. A total of 17 teachers participated with their classes in the testing.

The data obtained are primarily suitable for analysing specific usage patterns of the app among students from classes that demonstrably worked with the app. An important step before the analysis itself was determining an appropriate method to narrow the observed sample to a set of reliable users.

3.1 Data Selection

For the analysis, four courses (excluding summary quizzes) were selected: *Statistics* (C1); *Measurement in Geometry* (C2); *Fractions* (C3); *Equations and inequations* (C4). It was decided to analyse topics that differ not only from a mathematical point of view (statistics, geometry, arithmetic and algebra) but also in their design (different numbers of blocks, block lengths, and types).

To better analyse students' behaviour across lessons, the lesson blocks were coded. The coding was done by two researchers and compared. Based on the coding, six broader categories (block types) emerged (see Table 1). These were further refined and discussed until a consensus among coders emerged. The *Brief non-explanatory instruction* (B1) block provides short information or text, such as "Let's try it!" and does not offer any additional learning information. A *Short text* (B2) is used as motivation, a lesson introduction, or a recapitulation of the lesson or its content part. *Explanatory text* is the main matter of the lessons in which students are expected to learn or revise the matter. The explanatory text blocks varied greatly in length, so it was decided to distinguish between shorter (B3) and longer (B4) ones using two categories. The other block type in which learning is expected is the block that includes an interactive task. We distinguish between the *Basic task* (B5) and the *Advanced task* (B6).

Table 1: The number of block types in courses

Course / Block Type	B1	B2	B3	B4	B5	B6	Total
C1	0	2	8	7	1	1	19
C2	0	4	4	2	3	4	17
C3	6	4	9	18	7	3	47
C4	12	7	11	20	7	6	63
Total	18	17	32	47	18	14	

The *Statistics* course (C1) was primarily lecture-based and consisted of two lessons. Courses C3 and C4 were similar in structure, containing 3 and 5 lessons, respectively. In both courses, the percentage of exercises was around 20%; overall, *Longer explanatory blocks* (B4) predominated, but there was also a significant proportion of B1 blocks, i.e., blocks not expected to require high cognitive or time demands. Course C2, on the other hand, was more task-oriented (blocks with tasks accounted for 41.2% of the course), and the content blocks tended to be concise.

3.1.1 Narrowing the Sample

For this paper, it was necessary to narrow the sample. Data reduction involved selecting classes with a certain minimum number of students participating in the selected courses. The goal was to include entire groups of students (classes) rather than just individuals, to avoid bias arising from selecting only active participants. In this context, the following criteria were established to identify the selected classes of students: (1) a minimum average of 13 students per class across all 4 courses, (2) a minimum of 9 students in each course, and (3) a minimum average of 33 interactions per student in the given class.

Based on these criteria, 9 classes (G1–G9) were selected, comprising 201 unique students. A more detailed breakdown of students by class and participation in individual courses is shown in Table 2. The total number of students in each course does not add up to 201, because some students in each class did not complete all the courses.

Table 2: Number of students and interactions in individual classes across courses

Class	C1		C2		C3		C4		All Courses	
	Students (n)	Student inter. (n)	Students (n)	Student inter. (n)	Students (n)	Student inter. (n)	Students (n)	Student inter. (n)	Students (Mean)	Student inter. (Mean)
G1	19	348	19	358	19	864	18	1175	18.8	36.6
G2	21	385	22	409	20	953	23	1454	21.5	37.2
G3	12	228	19	339	13	611	22	1372	16.5	38.6
G4	13	235	10	180	13	611	17	985	13.3	37.9
G5	15	268	14	235	14	613	21	1258	16.0	37.1
G6	11	184	18	318	17	690	15	875	15.3	33.9
G7	16	297	17	306	15	667	20	1205	17.0	36.4
G8	18	342	21	378	20	907	22	1386	20.3	37.2

Class	C1		C2		C3		C4		All Courses	
	Students (n)	Student inter. (n)	Students (n)	Student inter. (n)	Students (n)	Student inter. (n)	Students (n)	Student inter. (n)	Students (Mean)	Student inter. (Mean)
G9	9	171	20	359	21	987	22	1393	18.0	40.4
Total	134	2458	160	2882	152	6903	180	11103		

3.2 Data Analysis Procedure

The analysis evaluated the course design and its individual blocks using interaction data collected by the Edu-AI application. We decided not to automatically remove outliers from the time data captured in the application logs. Both extremely short and extremely long times may indicate the nature of user interaction with a given block, particularly regarding gaming behaviour, inactivity, or technical interruptions. At the same time, the way response-time outliers are handled can significantly influence the interpretation of results (Kovanović et al., 2015; Vankov, 2023). Therefore, user interaction times were classified rather than simply removed.

Interactions were classified separately for each block into four duration-based categories: *short_irrelevant*, *short_relevant*, *long_relevant*, and *long_irrelevant* as described below.

The classification combined relative and absolute thresholds, drawing on the MAD-based approach described by Leys et al. (2013) and the time-on-task issues discussed by Kovanović et al. (2015). Because the observed interaction-time distributions were strongly right-skewed, times were log-transformed before the median absolute deviation (MAD) was calculated. The MAD was multiplied by 1.4826, the conventional consistency factor that makes it comparable to the standard deviation under normality. The main threshold was set at 2.5 scaled MAD from the block-specific median. Although more conservative thresholds, such as 3 MAD, have also been proposed (Leys et al., 2019), the more sensitive threshold was retained because the aim was to classify potentially informative interactions rather than exclude them.

For each block, the lower and upper thresholds were calculated as the block-specific median plus and minus 2.5× scaled MAD, respectively, on the log-transformed scale, and then transformed back into seconds. If the lower threshold was too small (less than 1 s), it was set to 1 s. Similarly, when the upper threshold was below 600 s, it was treated as 600 s. Interactions at or below the lower threshold were classified as *short_irrelevant*; those above the lower threshold and at or below the median as *short_relevant*; those above the median and at or below the upper threshold as *long_relevant*; and those above the upper threshold as *long_irrelevant*.

The categories were retained as indicators of potentially different interaction patterns, not as direct measures of students' attention, engagement, or understanding.

4. Results

Overall response-time statistics for individual blocks per course are presented in Figure 2. It presents the mean of the medians of interaction time for each block type in the specific course. The number of interactions per block ranged from 124 to 180, with an average of 153.

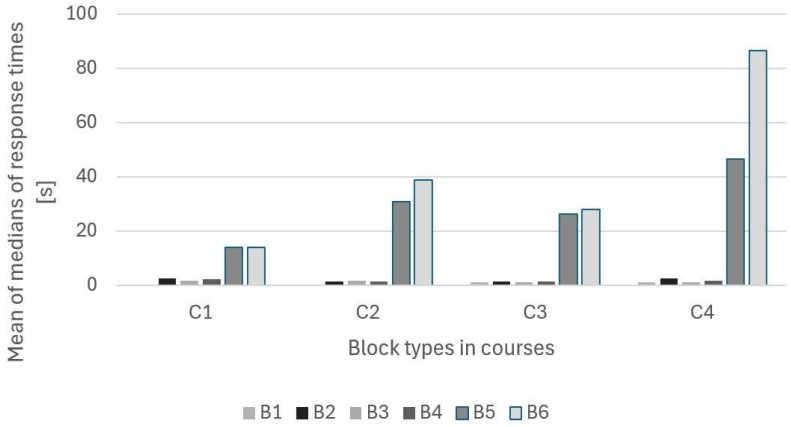


Figure 2: Mean of medians of response times (MMRT) for all interactions in courses and blocks by their type.

Response times were extremely low for blocks B1–B4. The block-level medians, summarised as the mean of medians, indicate that a typical block of this type was passed very quickly, often within one or two seconds. This suggests that many students did not spend enough time on explanatory or introductory blocks for substantial engagement with the content to be likely. For example, in the *Statistics* course, the median interaction time in blocks classified as longer explanatory text (B4) was very low, with block-level medians ranging only from approximately 1 to 4 seconds. This suggests that many student-block interactions with explanatory content were very brief.

If these blocks are compared with the task blocks (B5, B6), a surge in response times can be observed. For example, in the C3 course *Fractions*, 7 blocks were classified as *Basic tasks* (B5). In these, roughly half of the students spent more than 26 seconds on average. While this can be partially explained by technical differences, it is not necessarily the case. Figure 2 also shows that the medians distinguished between basic and advanced tasks (as classified by the authors). In the case of the C4 course, *Equations and inequations*, this difference is striking, indicating the actual attention paid to the tasks.

However, given the very steeply skewed log-normal distribution, this overview has its visible limitations. Response time medians provide a good estimate of whether gaming occurs in the majority of the sample and also help identify significant systemic differences (such as those between tasks and other blocks). On the other hand, they do not provide further information, for example, on the specific numbers of interactions and their characteristics. We attempted to capture these using the analysis described in section 3.2.

4.1 Evaluation of Block Types Based on the Student-Block Interactions

The MAD-based classification provided only limited additional information as a four-category classifier, especially for explanatory blocks. This limitation follows from the extremely short interaction times observed, mostly in non-task blocks (B1-B4). When a substantial part of the sample spends only one or two seconds on a block, the distinction between *short_irrelevant* and *short_relevant* becomes difficult to interpret substantively. In this context, the block-specific median appears to be the most practically meaningful threshold. Interactions below or at the median indicate only minimal engagement with the block in the learning context. This phenomenon is evident for all information blocks (B1 - B4) regardless of their scope and purpose (explanation of the material, motivation, recapitulation, etc.). The results show us that for these block types in most cases the interactions were so short that the students could not even read the content or cognitively grasp it, especially for explanatory blocks (B3, B4).

In the case of blocks B5 and B6, our method of response-time classification provides more space for reasonable interpretations. Within block B5, we identified a not-negligible number of relevant interactions (see Table 3), The representation of relevant interactions across all courses was 96%, with the lowest value of 95% in C4. It can be observed that the relative proportion of individual interaction types (short/long relevant/irrelevant) does not significantly correlate with the mean of medians (MMRT) in individual courses (see Figure 2), and is therefore essentially independent of the length of solving tasks (even without counting C1, where there was only one B5 block).

Table 3: Overview of interactions for blocks B5

Course	Number of Blocks (n)	Number of Interactions (n)				
		Total	short_irrelevant	short_relevant	long_relevant	long_irrelevant
C1	1	132	0	71	60	1
C2	3	476	0	240	232	4
C3	7	1038	40	465	532	1
C4	7	1233	51	550	624	8

5. Discussion

All recorded interaction times were directly related to the phenomena under investigation, including rapid progression through the app, possible gaming behaviour, inactivity, and technical interruption. Therefore, instead of trimming short and long outliers, we used a method (Leys et al., 2013) to classify student-block

interactions by their duration. Other methods for defining response-time thresholds (Stipek et al., 2025) could also be applied and may capture different aspects of student behaviour. Further work should therefore compare alternative classification approaches and examine their suitability for this type of application data.

At this stage, the data allow only a tentative comparison between courses, but they reveal interesting findings regarding block types. In longer explanatory blocks (B4), the share of *long_relevant* interactions was higher in the shorter courses C1 and C2 than in C3, and C4 occupied an intermediate position. This pattern suggests a working hypothesis that course length and the number of explanatory blocks may influence students' willingness to engage with explanatory content. We also observe a significant difference in interaction times between content-type blocks (B1-B4) and task blocks (B5, B6). For blocks B1-B4, we observe very short interactions, indicating little interest in the content presented. This lack of interest may be due to the fact that students in this type of application are more interested in solving tasks (e.g. similar to Duolingo). From a gamification perspective, the cause can also be identified at the motivational level, where a higher cognitive grasp of the content does not bring any benefit compared to quick skimming.

If any blocks captured students' attention, they were the tasks (B5 or B6). The sample selection might have influenced this, as students were just a few weeks before the entrance examination, and those who had been preparing themselves preferred to revise tasks similar to the examination format. In these blocks, we can observe more time-meaningful interactions, with a lower level of gaming and a more expected effective impact on student learning. However, it is necessary to take into account secondary factors that may influence the interaction time and that we are not able to detect using the method employed (e.g., viewing the solution or previous blocks). The four-item structuring of interactions also appears to be beneficial for subsequent, more detailed data analyses. For instance, in course C3, we can see 40 *short_irrelevant* interactions within B5, which can be used to identify ill-behaved tasks (which are, for example, gamed more often than others) or to model student behaviour (for example, identifying those that have not one but several irrelevant interactions).

In terms of block classification, the method used also more or less confirmed the correct classification of Basic and Advanced task blocks, with the difference in time spent on these blocks being obvious, especially in course C4 (see Figure 2).

6. Conclusion

The present analysis suggests two follow-up tasks. First, it is necessary to distinguish students whose interactions fall into irrelevant categories but who may nevertheless have completed the course responsibly, for example, because they processed some blocks quickly but consistently. Second, it is necessary to identify students whose overall interaction time exceeds the median yet still exhibit local gaming behaviour in specific parts of the course. These two groups are important because they cannot be reliably classified by overall course time alone. Together, they form a suitable basis for a secondary analysis aimed at evaluating individual blocks more precisely: not only whether a block was passed quickly or slowly, but whether its position, format, and function in the course contributed to meaningful engagement or encouraged superficial progression through the app.

In relation to the course or lesson structure, based on the findings we shall consider including more task type blocks that seem more engaging for students, or to transform explanatory blocks into another type that requires more interaction or connects more with the task type block.

In the future, it is also important to incorporate more features that limit the gaming phenomenon in the app, including testing new, more engaging block types. Additional qualitative data (such as interviews with students or teachers' feedback) are needed to help clarify some questions about blocks, students' behaviour in the app, and other observed phenomena.

Acknowledgements

This paper was supported by the project TAČR N. TQ23000334 – AI-based Adaptive Learning Platform.

Ethics Declaration

The research did not involve any personally identifiable information about the students. The data was processed in an anonymised form.

AI Declaration

Generative AI tools were used in the preparation of this article for graphic design, language support, and text organisation. The authors of this article developed all technical concepts, methodologies, and interpretations.

References

Arroyo, I., Ferguson, K., Johns, J., Dragon, T., Meheranian, H., Fisher, D., Barto, A., Mahadevan, S., & Woolf, B. P. (2007). Repairing disengagement with non-invasive interventions. In *Proceedings of the 13th International Conference on Artificial Intelligence in Education* (pp. 195–202). IOS Press.

Baker, R. S. J. d., Corbett, A. T., Koedinger, K. R., Evenson, S., Roll, I., Wagner, A. Z., Naim, M., Raspat, J., Baker, D. J., & Beck, J. E. (2006). Adapting to when students game an intelligent tutoring system. In M. Ikeda, K. D. Ashley, & T.-W. Chan (Eds.), *Intelligent tutoring systems: ITS 2006* (Lecture Notes in Computer Science, Vol. 4053). Springer. https://doi.org/10.1007/11774303_39

Canilao, C. M., & Gurat, M. G. (2023). Effect of Mobile App on Students' Mathematics and Technology Attitude. *American Journal of Educational Research*, 11(10), 722–728.

Cates, S., Barron, D., & Ruddiman, P. (2017). MobiLearn Go: Mobile microlearning as an active, location-aware game. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Article 103, pp. 1–7). Association for Computing Machinery. https://doi.org/10.1145/3098279.3122146

Guill, K., Lüdtke, O., & Köller, O. (2020). Assessing the instructional quality of private tutoring and its effects on student outcomes: Analyses from the German National Educational Panel Study. *British Journal of Educational Psychology*, 90(2), 282–300. https://doi.org/10.1111/bjep.12281

Jančařík, A., Michal, J., & Novotná, J. (2023). Using AI Chatbot for Math Tutoring. *Journal of Education Culture and Society*, 14(2), 285–296.https://doi.org/10.15503/jecs2023.2.285.296

Kadhem, H. (2017). Using mobile-based micro-learning to enhance students' retention of IT concepts and skills. 2017 International Conference on Knowledge Engineering and Applications (ICKEA), 128–132. https://doi.org/10.1109/ICKEA.2017.8169915

Kapp, K. M., & DeFelice, R. A. (2019). *Microlearning: short and sweet*. ATD Press.

Kovanovic, V., Gašević, D., Dawson, S., Joksimovic, S., & Baker, R. (2015). Does time-on-task estimation matter? Implications on validity of learning analytics findings. *Journal of Learning Analytics*, 2(3), 81–110.https://doi.org/10.18608/jla.2015.23.6

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4), 764–766. https://doi.org/10.1016/j.jesp.2013.03.013

Mabhamba, M., Mukuka, A., & Alex, J. K. (2024). Evaluating the effectiveness of private supplementary tutoring on grade 12 learners' mathematics achievement. *Eurasia Journal of Mathematics, Science and Technology Education*, 20(9), em2507. https://doi.org/10.29333/ejmste/15168

MŠMT. (2023). *Rámcový vzdělávací program pro základní vzdělávání*. Výzkumný ústav pedagogický. https://www.edu.cz/wp-content/uploads/2023/07/RVP_ZV_2023_cista_verze.pdf

Průcha, J. (2015). *Česká vzdělanost: Multidisciplinární pohled na fenomén národní kultury*. Praha: Wolters Kluwer.

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

Şahin, Z. G., & Kirmizigül, H. (2026). Reflections from the mathematics lesson with micro learning: A different educational experience in the digital age. J*ournal of Learning and Teaching in Digital Age*. Advance online publication. https://doi.org/10.53850/joltida.1702683

Son, T. (2024). Intelligent Tutoring Systems in Mathematics Education: A Systematic Literature Review Using the Substitution, Augmentation, Modification, Redefinition Model. *Computers*, 13(10), 270. https://doi.org/10.3390/computers13100270

Štípek, J., Voňková, H., & Králová, K. (2025). Vztah rychlého odpovídání v testu z anglického jazyka a výsledného skóre u různých skupin žáků. *Orbis Scholae*, 19 (1), 41-66. https://doi.org/10.14712/23363177.2025.7

Šťastný, V. (2016). *Fenomén soukromého doučování jako stínový vzdělávací systém v České republice* [dizertační práce]. Univerzita Karlova, Pedagogická fakulta, Oddělení pro vědeckou činnost.

Taleb, Z., Ahmadi, A., & Musavi, M. (2015). The effect of M-learning on mathematics learning. *Procedia - Social and Behavioral Sciences*, 171, 83–89.https://doi.org/10.1016/j.sbspro.2015.01.092

Tang, J., Wijaya, T. T., Weinhandl, R., Houghton, T., Lavicza, Z., & Habibi, A. (2022). Effects of Micro-Lectures on Junior High School Students' Achievements and Learning Satisfaction in Mathematics Lessons. *Mathematics*, 10(16), 2973. https://doi.org/10.3390/math10162973

Vankov, I. I. (2023). The hazards of dealing with response time outliers. *Frontiers in Psychology*, 14, 1220281.

Zhang, Y., Cui, C., He, Y., & Wang, L. (2022). Does private supplementary tutoring matter in Chinese students' learning of mathematics: A longitudinal study. *ZDM – Mathematics Education*, 54(3), 737–747. https://doi.org/10.1007/s11858-022-01346-6