Computers and Education: Artificial Intelligence 9 (2025) 100483

Contents lists available at ScienceDirect

Computers and Education: Artificial Intelligence

journal homepage: www.sciencedirect.com/journal/computers-and-education-artificial-intelligence

Artificial intelligence in classroom management: A systematic review on educational purposes, technical implementations, and ethical considerations

Tim Fütterer [a,*], Patricia Goldberg [a], Babette Bühler [a,b], Vlasta Sikimić [c], Ulrich Trautwein [a], Peter Gerjets [d], Kathleen Stürmer [a], Enkelejda Kasneci [b]

[a] *University of Tübingen, Tübingen, Germany*
[b] *Technical University of Munich, Munich, Germany*
[c] *Eindhoven University of Technology, Eindhoven, Netherlands*
[d] *Leibniz-Institut für Wissensmedien, Tübingen, Germany*

A R T I C L E I N F O

Keywords:
Artificial intelligence
Classroom management
Attention
Ethics
Multimodal data

A B S T R A C T

Artificial intelligence (AI) is increasingly shaping education, including classroom management. This systematic review analyzes 104 studies (2000–2022) on the use of AI in classroom management, focusing on educational purposes, technical implementations, and ethical considerations. Our findings show a growing use of AI technologies—particularly machine learning and deep learning—for tasks such as attendance tracking, behavior monitoring, and engagement assessment. These tools can streamline classroom management and offer detailed insights into student behavior. However, only a minority of studies leveraged AI's full potential, such as real-time feedback or multimodal data. Ethical issues, particularly privacy, data security, and algorithmic bias, were often underreported: only 22 % of studies addressed ethical concerns, and just 13 % implemented privacy-preserving measures. Our review underscores the importance of balancing technological innovation with ethical responsibility. It offers a comprehensive overview of AI's current applications and highlights future challenges and directions for responsible AI use in classrooms.

1. Introduction

Efforts in educational research and computer sciences are increasing worldwide to develop artificial intelligence (AI) systems that can capture lesson activities (e.g., classroom management) automatically (e.g., Foster et al., 2024; Hou et al., 2024), no doubt fueled by the rapid technological developments in (generative) AI. As such, AI systems are becoming a reality in school practice, so it is important to understand for which educational purposes such technology is (or could be) used and what role ethical aspects play. For instance, in 2019, the Wall Street Journal published a video (youtube.com/watch?v=JMLsHI8aV0g) reporting that a growing number of classrooms in China are equipped with cameras and brain-wave trackers that allow teachers and parents to keep track of their students'. By the end of August 2024, the video had more than 3.5 million views and attracted more than 12,000 comments on the YouTube platform, most of them highly critical, with one contributor calling the application of this kind of AI in classrooms a *mental prison* and many contributors calling for strictly restraining the use of AI in education. The Wall Street Journal's report on AI technologies in Chinese schools is an excellent example illustrating how AI is already used in schools. On the one hand, such examples promise that AI could be used to organize teaching more effectively by simplifying and accelerating the automation of monitoring activities and, potentially, facilitating adaptive teaching practices that consider each student's level of knowledge and interest. AI's efficiency can analyze vast amounts of data at a speed humans cannot match. On the other hand, the same AI can be used to control students, especially if it is questionable how valid the automated assessment is and to what extent the consequences for learners are permissible. The video illustrates that great care is needed when using AI in education—with ethical considerations being the highest concern. Against this backdrop, the role of AI in education becomes a double-edged sword, promising advances in efficiency and effectiveness but posing important challenges to privacy and ethical standards in education.

* Corresponding author. Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Europastraße 6, 72072, Tübingen, Germany.
E-mail address: tim.fuetterer@uni-tuebingen.de (T. Fütterer).

https://doi.org/10.1016/j.caeai.2025.100483
Received 25 December 2024; Received in revised form 13 August 2025; Accepted 24 September 2025
Available online 25 September 2025

Therefore, it is necessary to step back and clarify several issues: To what extent does this perhaps best-known example of using AI in education for classroom management reflect current research approaches and outcomes on this topic? How much research has been done on the topic, what kind of research has been done, and for what purposes? What technologies have been used so far? And what is the role of ethical considerations in current contributions? We conducted a systematic literature review focusing on the AI applications developed for use (i.e., are already useable) in classrooms restricted to classroom management. Classroom management is considered a crucial aspect of effective teaching and is characterized, for instance, by providing structure for students, such as rules and procedures, or maintaining students' attention to lessons and their engagement in activities (Patall et al., 2023; Seidel & Shavelson, 2007). Incorporating AI into classroom management is of paramount interest when it comes to improving learning as it can enhance the learning environment by automating administrative tasks, allowing teachers to devote more time and attention to interactive and personalized teaching rather than being preoccupied with routine management tasks (especially those tasks rather distant from teaching content such as tracking the attendance of students). Moreover, automation of administrative tasks is becoming more important considering the challenge of teacher shortages in many countries.

2. Theoretical background

2.1. Classroom management as a key characteristic of effective teaching

Classroom management plays a crucial role in students' success in learning. In teaching effectiveness research, there is a consensus that teachers successfully support student learning when they organize effective teaching (for teachers' classroom management practices such as providing structure, see Patall et al., 2023). Effective teaching is characterized, among other things, by monitoring student behavior, ensuring clarity and transparency about goals and rules in the classroom, providing a structure that guides students' behavior, and effectively managing disruptions (i.e., increasing students' active learning time; Brophy, 1983; Praetorius et al., 2018; Seidel & Shavelson, 2007).

Such measures to organize and guide students' learning (e.g., providing structure or effectively managing disruptions) can be subsumed under a teaching effectiveness dimension called *classroom management* (Emmer & Evertson, 1981). Following Brophy (2006), classroom management can be defined as "actions taken to create and maintain a learning environment conducive to successful instruction (arranging the physical environment, establishing rules and procedures, maintaining students' attention to lessons and engagement in activities)" (p. 17).

Classroom management leads to a higher probability of deeper processing by students, more effective use of time, and more positive emotions in students during learning (Klieme et al., 2009). For instance, based on synthesizing analyses of 179 handbook chapters and reviews and 91 meta-analyses, Wang et al. (1993) found that classroom management is one of the most influential factors leading to increased student engagement, decreased disruptive student behavior, and less time wasted during lessons, positively impacting students' academic achievement. Consistently, in his synthesis of more than 800 meta-analyses, Hattie (2009) showed that maximizing learning time (i.e., time on task) has an effect of $d = 0.38$ on student learning performance, and Kunter et al. (2007) found that students individually perceived classroom management strategies are associated with developing subject-related interests. Furthermore, effective classroom management seems especially important in elementary education (e.g., motivational-affective outcomes: Seidel & Shavelson, 2007). Correspondingly, Korpershoek et al. (2016) found an effect of $g = 0.22$ on students' academic, behavioral, and social-emotional outcomes.

Classroom management represents a requirement for students to become attentive and an indicator of students' actual attention (Praetorius et al., 2018). Therefore, the success of classroom management always emerges from a complex interaction between teachers and students (Korpershoek et al., 2016). However, looking at the definitions of classroom management—as the one above—teachers play a central role in classroom management. Teachers face the challenge of using different strategies (e.g., explaining rules of conduct before disturbances occur in the spirit of cautious-restrained handling of disturbances or a "low-profile approach": Kounin, 1970) to organize students' learning in the classroom effectively. The importance of teachers in classroom management is also evident in the description of core teaching practices, many of which relate to classroom organization (i.e., implementing organizational routines: Grossman, 2018).

As classroom management activities take up teachers' time and cognitive capacities, which could be used more sensibly (e.g., to support weaker or stronger learners), and because some of these activities (e.g., monitoring students) can be automated and potentially analyzed at a more granular level, there is hope that AI-based applications for classroom management can relieve teachers (e.g., reduction of cognitive load). In addition, there is justified hope in educational research and psychology that more reliable and valid AI systems can replace (sometimes unreliable) external observers.

2.2. Artificial intelligence in education

2.2.1. Opportunities and risks of artificial intelligence in classroom management

AI applications in education, and particularly in classroom management, offer multiple supportive opportunities. A fundamental aspect of effective classroom management, as highlighted by Grossman (2018), Brophy (1983), and Kounin (1970), involves teachers continuously monitoring their students and the dynamics of student interactions. This task, while crucial, demands a significant investment of time and attention, a resource increasingly scarce amidst growing global teacher shortages. Whereas dedicating time to classroom-related tasks is essential for teaching efficacy, many subtasks, such as monitoring student attendance, represent an inefficient use of teachers' time. This is underscored by findings from a representative survey of 1,032 teachers at general and vocational schools in Germany (Robert Bosch Stiftung, 2023), which identify these tasks as potential time wasters.

In this context, AI systems can be supportive, not merely automating tasks but enhancing the teaching experience. By handling demanding and time-consuming tasks, such as tracking attendance in large classrooms, AI can free teachers to focus more on pedagogical aspects. Moreover, as suggested by Korpershoek et al. (2016) and Lewis and Sugai (1999), effective classroom management leans towards proactive techniques over reactive ones. AI technology can bolster this proactive approach, for instance, through alert systems that preempt issues rather than serve as a basis for sanctions.

Additionally, AI can aid in fostering a more objective classroom environment, helping teachers to engage in critical self-reflection. By providing data-driven insights, AI tools can reveal underlying patterns, including potential implicit biases related to gender, race, or other factors (Deutscher Ethikrat, 2023), thus contributing to a more equitable educational setting.

The tangible experience of support and the noticeable reduction in workload that AI can bring are vital to increasing teachers' acceptance of these technologies. Such acceptance is crucial for AI's successful, widespread integration in classrooms. By positioning AI as a supportive tool rather than a mere automaton, we can better harness its potential to enrich the educational landscape, benefiting teachers and students.

In addition, regarding educational research, the valid and reliable measurement of classroom management constitutes a challenge for researchers. Traditional classroom management approaches often operate on high-inference levels, making them prone to subjective biases. Traditional approaches also reach their limits, especially when aiming at real-time measurements. Neither classroom observations by humans nor teachers' and students' self-reports can provide the kind of information technology can offer (see the explanations in the discussion section). Visual, auditive, or physical (e.g., accelerometer) signal analyses have the potential to (a) collect more information simultaneously and continuously (e.g., of each student at the same time for the whole lesson), (b) to do so unobtrusively without the need to interrupt the course of action for filling in questionnaires et cetera, and (c) to make this information directly available so that it can be exploited for adaptive interventions. As some indicators of successful classroom management are directly observable, technology can be deployed to collect low-inference data to identify, for example, disruptive behavior or efficient use of instructional time (Emmer & Evertson, 1981). Researchers may gather data on student behavior, such as attendance rates, disciplinary actions, and academic performance, as well-managed classrooms are often associated with fewer discipline problems and improved student engagement and achievement. Furthermore, standardized tools can assess the classroom's physical environment and how it supports effective management. This can include factors such as seating arrangements, classroom organization, and the presence of learning materials.

With the progress of technology, however, AI presents itself not only with promising potential for practice and research but also with some undeniable risks. Especially in education, the responsible use of AI technology is important. According to the OECD, properly functioning AI systems (i.e., they do what they are supposed to) are as important as humans' socially just, fair, and non-discriminant usage (Vincent-Lancrin & van der Vlies, 2020). In this vein, critics consider current technology insufficient to analyze attention and emotions in the classroom reliably and validly (Goldberg et al., 2021) and worry about improper interpretation and usage of student data (e.g., Deutscher Ethikrat, 2023; Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence [Artificial Intelligence Act| and Amending Certain Union Legislative Acts, 2021). Thus, the risks of AI in classroom management can be manifold. For instance, researchers and developers need to ensure that the technology functions objectively, reliably, and validly. Moreover, if not implemented properly, attention monitoring or affect recognition can create an environment focusing on achievement rather than learning, leading to—at least a feeling of—surveillance and constant pressure. Moment-to-moment information about students' state of mind or their progress in completing tasks enables direct sanctioning or conditioning, and learners might experience an invasive intrusion into their autonomy and privacy (Deutscher Ethikrat, 2023), which is heightened if the gathered information leaves the safe space of the classroom and is communicated to higher authorities or parents. For instance, intrusion would be present in the use of intelligent tutoring systems that record data at the time students do homework. Panchal and Mohammad (2020) reported how AI is already implemented in Chinese schools. They explained that, in some cases, facial recognition and surveillance cameras are used to track students' attendance and location on the school campus and to record the number of yawns or how often they check their phones. Additionally, EEG headbands could be deployed to measure concentration and report this information to parents. However, not only may students and the learning atmosphere suffer from this kind of close monitoring, but also teachers may experience negative impacts on their self-perception (Deutscher Ethikrat, 2023). Additionally, teachers might feel insecure because of insufficient knowledge and understanding of AI systems or a lack of technological literacy and new design skills.

2.2.2. *An introduction to the technical functionality of AI*

There are a variety of definitions for the term *AI* in scientific discourse. Baker and Smith (2019) defined AI in an educational context as "computers which perform cognitive tasks, usually associated with human minds, particularly learning and problem solving" (p. 10). The term encompasses a range of technologies and approaches, including Machine Learning (ML), Computer Vision, and Natural Language Processing (NLP). As an overarching concept, ML refers to automatically finding patterns and rules in data and then using this to make predictions on new, similar data (Domingos, 2012). One branch of ML is deep learning, which encompasses powerful and complex algorithms such as artificial neural networks. Deep learning algorithms have gained traction over the last decade, as they are successfully employed for challenging and data-intense tasks such as image recognition. This, in turn, makes them a crucial component of other AI subfields defined by the task they are solving, such as Computer Vision and NLP. Moreover, advanced classification techniques developed in the field of bioinformatics (e.g., Khan, Noor, et al., 2025; Khan, Uddin, et al., 2025) demonstrate the potential of ML models to handle high-dimensional, noisy, and complex data—challenges that are increasingly relevant for educational AI systems analyzing multimodal classroom data. Insights from these studies (e.g., the use of transformer-based embeddings and SHAP-enhanced feature selection in bioinformatics) may inspire the development of robust classification pipelines for detecting student behaviors or engagement patterns in real-time classroom environments.

With the improvement and further spread of these new methods, they are increasingly used to support students, teachers, and administration in the educational domain (Baker & Smith, 2019). Zawacki-Richter et al. (2019) identified four main areas of application of AI in the educational context: *profiling and prediction* (i.e., predicting academic achievement of students), *assessment and evaluation* (i.e., evaluation of student engagement), *intelligent tutoring systems* (i.e., provision of automated feedback), as well as *adaptive systems and personalization* (i.e., providing personalized content). However, these four application areas do not have to be regarded as independent (research) fields but can also overlap, for example, when considering the use of AI in classroom management. Diverse subcommunities, such as learning analytics and NLP applications in education, investigate various dimensions within their research. This review aims to offer a comprehensive overview of these subcommunities, focusing on classroom management.

In sum, many AI technologies could support teachers' classroom management, provide students with the best possible learning conditions, and support research by creating and analyzing more complex data. However, success depends not only on the technology and its flawless and unbiased functioning but also on human interpretation and usage. In this vein, data security and privacy protection need to be discussed alongside other ethical considerations.

2.3. *Ethical considerations when using AI for educational purposes*

Education is an ethically sensitive area for implementing AI. It can direct people's professional lives and determine their future. For this reason, AI assessments and predictions of academic success have been classified as high-risk (e.g., Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence [Artificial Intelligence Act| and Amending Certain Union Legislative Acts, 2021). In addition, predictive algorithms can reinforce biases (2022), and algorithmic assessments might limit teachers' and students' autonomy in decision-making.

Data is the basis for using AI in the classroom, especially in classroom management activities. A differentiated approach allows for two fundamental perspectives here: On the one hand, big data (e.g., continuous gaze or body movements of all students in a class) provide AI systems with the basis for interpreting student behavior and supporting algorithmized classroom management situations. In addition, big data allows researchers to optimize algorithms, better understand complex classroom situations, and ultimately promote student learning in the classroom. If data are used for research purposes, principles of good science suggest that data should be stored for enough time and made publicly available. On the other hand, surveillance might become a problem if big data are available. Access to big data about students can suddenly become the groundwork for (possibly unintentionally or abusively) tracking them on even more occasions and denying them the privacy that previous generations enjoyed. In addition, the concept of *raw data* has been questioned in the philosophy of science (e.g., Leonelli, 2015). Data are usually generated, curated, and stored with a particular purpose in mind, and this purpose was evaluated when obtaining the ethical permit. The secondary use of data might be very different and requires a new ethical permit. In addition, informed consent obtained by the participants might not cover the data repurposing for new research questions, especially if the application used was unknown when obtaining the consent. Hence, the principle of open science must be balanced with privacy for every use case.

Against this backdrop, it can be argued that the use of AI in education practice and research must follow the highest ethical standards. The importance of accompanying measures (e.g., laws, ethical guidelines) for using AI in educational contexts is demonstrated by the fact that more and more AI and data-protection regulations are being implemented (see the international efforts around safety AI). China, for instance, has issued the earliest and most detailed AI legislation, albeit these regulations apply only to companies' AI systems but not to the government (Hutson, 2023). Furthermore, the European Commission issued the so-called *EU Artificial Intelligence Act* (https://artificialintelligenceact.eu) in 2021, a comprehensive law on AI regulation (Brakel & Uuk, n.d.; Hutson, 2023), the final version of which was ratified by the European Parliament and the European Council on December 9, 2023, and accepted in March 2024. In the United States, several discussions—driven primarily by leading technology providers such as OpenAI, Meta, and Google—have taken place. In 2022, the White House Office of Science and Technology Policy released a blueprint for an AI Bill of Rights. Whereas Hutson (2023) noted that, as of mid-2023, no federal AI laws or major data protection regulations had been implemented, the United States has since taken regulatory action through Executive Order 14110 on the "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence" (October 30, 2023). This Executive Order emphasizes the importance of privacy, fairness, and safety in AI applications, including those used in educational contexts (https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence).

Privacy is one key aspect of general ethical considerations besides beneficence, justice, or transparency (e.g., Hagendorff, 2022). Privacy considerations are also relevant when applied to the educational context and specifically to AI technologies that can be used for classroom management. On the one hand, AI-based technologies can provide learners with interesting (e.g., individualized) and stimulating learning materials (e.g., virtual reality that can be used to recreate historical events). Moreover, AI-based learning systems can provide learners with personalized instruction, tasks, materials, or feedback (i.e., technology-supported adaptive teaching; Hardy et al., 2019). On the other hand, educational data can be (mis)used to predict someone's academic performance and work success or for job placements, for which *privacy criteria* need to be followed. Data gathered through education often contain personal information, which is why the use of AI on children comes with even more ethical concerns than its application on adults. For instance, the long-term—potentially negative—effects of AI applications on children (e.g., their self-regulated skills, self-efficacy, motivation to learn) are hard to predict, and questions regarding who should have the moral right to consent are open.

2.4. Research questions

This study focuses on AI applications used or potentially available for classroom management processes (e.g., assessing students' attention). This study aims to outline and describe the usage of AI applications in classroom management (*relevance*), to provide an overview of existing AI applications (*resources*), and to identify potential ethical limitations and how they were addressed in the studies (*responsibility*). To achieve these aims, we concentrated on the following three research questions (RQs):

(RQ1) What are the educational purposes of AI applications for classroom management?
(RQ2) What kind of AI applications are used or potentially available for classroom management (i.e., what is currently possible from a technical point of view)?
(RQ3) What are the ethical considerations regarding AI applications for classroom management?

3. Method

3.1. Systematic literature search

The systematic literature search followed the PRISMA guidelines (Page et al., 2021). The literature search utilized the Web of Science, PsycINFO, and Academic Search Premier databases in March 2022. We used these three databases to ensure broad access to publications in the field of psychology and education. We searched for studies that were published online between 2000 and March 2022. We chose this period because, in the last two decades, there has been a leap in the field of ML methods, especially in deep learning and computer vision, which can be confirmed with literature, especially from 2007 onwards. Moreover, we deliberately restricted the scope of this review to studies published up until March 2022 as this cut-off allows us to analyze a coherent body of literature representing the pre-generative AI (pre-genAI) era. Whereas we acknowledge the importance of recent advances in genAI, including tools such as ChatGPT, we argue that combining pre- and post-genAI studies within a single review risks conflating fundamentally different technological paradigms. By focusing on the pre-genAI period, we are able to provide a systematic synthesis of classroom management applications grounded in conventional machine learning and computer vision technologies, offering a baseline against which future genAI-focused research can be compared. Our search string included the three content areas: *teaching*, *artificial intelligence*, and *classroom management* (see Table 1 for the search string in detail). The literature search resulted in a total of 5,062 studies. Duplicates were automatically identified and deleted based on DOI and authors, title, and year in an automated process utilizing *R* (R Core Team, 2022) and *R* Studio version 2022.02.1 (Posit team, 2022) across the three databases and within each database. After removing 832 duplicates within and across databases, our initial data set for the screening phase of the systematic review included 4230 studies. The flow chart (Fig. 1) provides an overview of the systematic search and selection of studies.

Table 1
Taxonomy used for search in databases.

Topic	Search Terms
Learning setting	classroom* *OR* instruction* *OR* "virtual classroom*" *OR* "class session*"
AND	
Artificial intelligence	"artificial intelligence" *OR* "machine intelligence" *OR* "intelligent virtual realit*" *OR* "machine learning" *OR* "neural network*" *OR* "natural language processing" *OR* NLP *OR* "machine vision" *OR* "AI" *OR* ML *OR* "classroom analytics" *OR* "behavio?r analytics" *OR* "teaching analytics" *OR* "data science*" *OR* "computer vision" *OR* "deep learning"
AND	
Classroom management	observation* *OR* orchestration *OR* management *OR* monitoring *OR* awareness OR behavio?r *OR* engagement *OR* attention* *OR* motivation* *OR* emotion* *OR* learn* *OR* teacher* *OR* "teacher training" *OR* "teacher professionali?ation" *OR* "professional development" *OR* PD *OR* dashboard* *OR* classware *OR* "real?time" *OR* "professional vision" *OR* learn*

Note. This search string was adjusted to specific requirements of databases (e.g., symbols and function of wildcards).

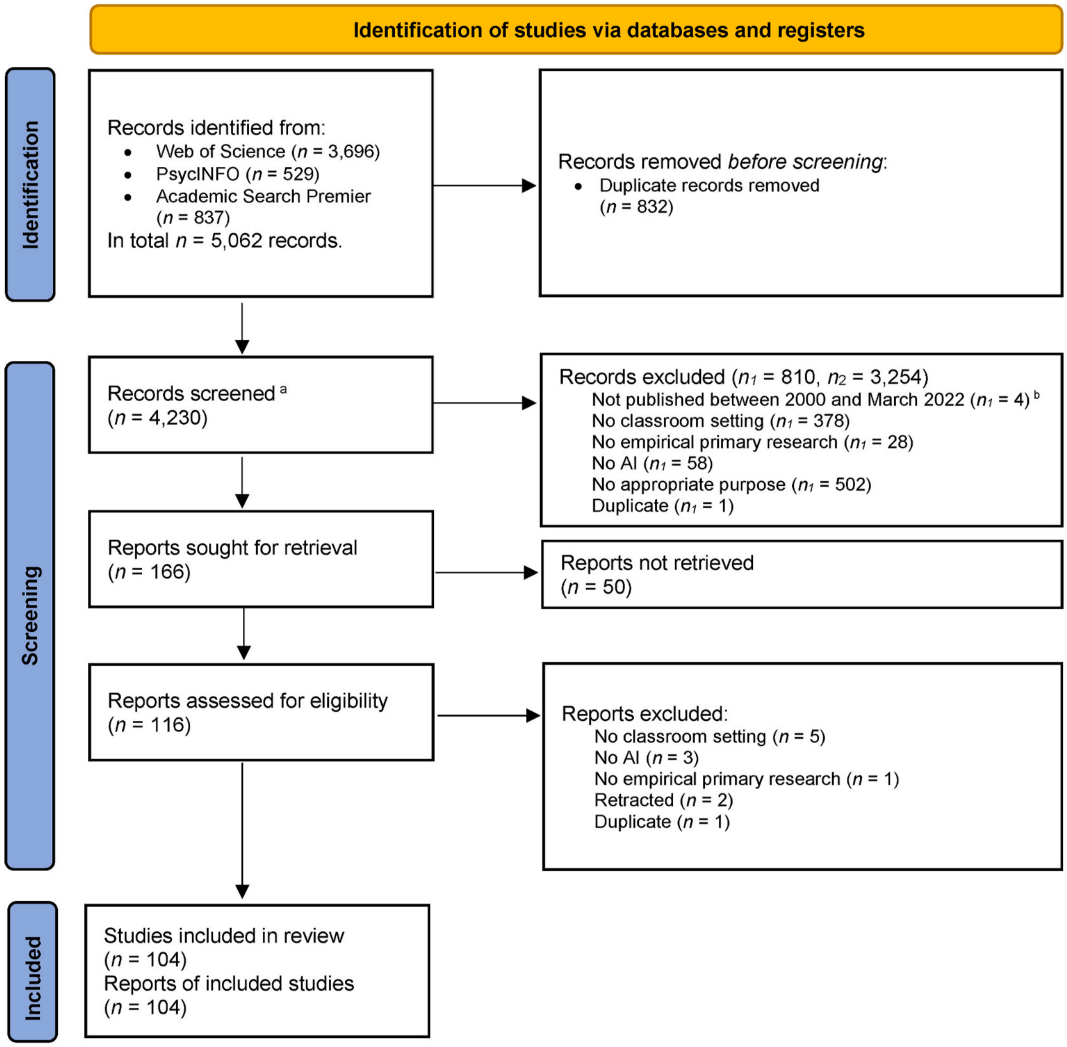

Fig. 1. Flow Diagram of Screening Process
Note. This flowchart follows the specifications according to PRISMA Statement. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021; 372:n71. doi: 10.1136/bmj.n71.
AI = artificial intelligence.
[a] Partly AI-assisted.
[a] Based on the 20 % random sample of records screened without artificial intelligence. n_1 = excluded records without using the artificial intelligence tool ASReview to assist the screening process, n_2 = excluded records when using the tool ASReview to assist the screening. Multiple reasons for exclusion per study were allowed.

3.2. Screening strategies: researcher-in-the-loop machine learning

We performed a two-step AI-assisted (i.e., active learning) screening process to filter relevant studies from the initial data set utilizing the tool ASReview (van de Schoot et al., 2021). More specifically, we used a combination of human ratings and AI support for screening—called *Human-in-the-Loop machine learning* (Holzinger, 2016) or *Researcher-in-the-Loop machine learning* for systematic reviews (Van de Schoot & De Bruin, 2020, Fig. 2)—to ensure an optimal cost-benefit ratio.

First, we trained two raters intensively (4 h) to independently screen n = 846 randomly selected studies (20 %) of the 4230 studies in the initial data set based on study titles and abstracts using inclusion criteria (criteria 1 to 6 in Table 2). The inter-rater reliability was Kappa (k) = 0.50, which can be classified as moderate (Landis & Koch, 1977). The first two authors made a final decision (consensus validation) for all studies where the two raters disagreed. Thirty-six studies were identified to fulfill all the inclusion criteria.

Second, we used two approaches to train the active learning model. In the first approach, we used the default setting in ASReview, which usually has a fast and excellent performance (van de Schoot et al., 2021). We used the feature extraction technique *term frequency-inverse document frequency* (TF-IDF), the classifier *Naive Bayes*, the query strategy *maximum*, and the balance strategy *dynamic resampling (double)*. In the second approach, we used the feature extraction technique (Reimers & Gurevych, 2019) in combination with the classifier *random forest,* as it was shown that these settings are particularly suitable for studies in the domains of education/educational psychology/psychology (Campos et al., 2024). These settings are suitable because greater consideration is given to semantics in abstracts (i.e., context), which is particularly advantageous for ambiguous constructs (i.e., no linkage to univocal terms) such as those commonly found in education and psychology. For both approaches, we used the already identified 36 studies plus five additional pre-identified key papers (i.e., studies that meet all inclusion criteria) and all 810 excluded studies as prior knowledge for the active learning model. The first author of this study performed the AI-assisted screening using inclusion and exclusion criteria (criteria 1 to 6 in Table 2). As recommended, the first authors stopped screening after 50 studies in a row were excluded, and at least 25 % of all 4230 studies were screened (Campos et al., 2024). In total, we identified 166 studies that fulfill all inclusion criteria. If the full text was not available, we asked the

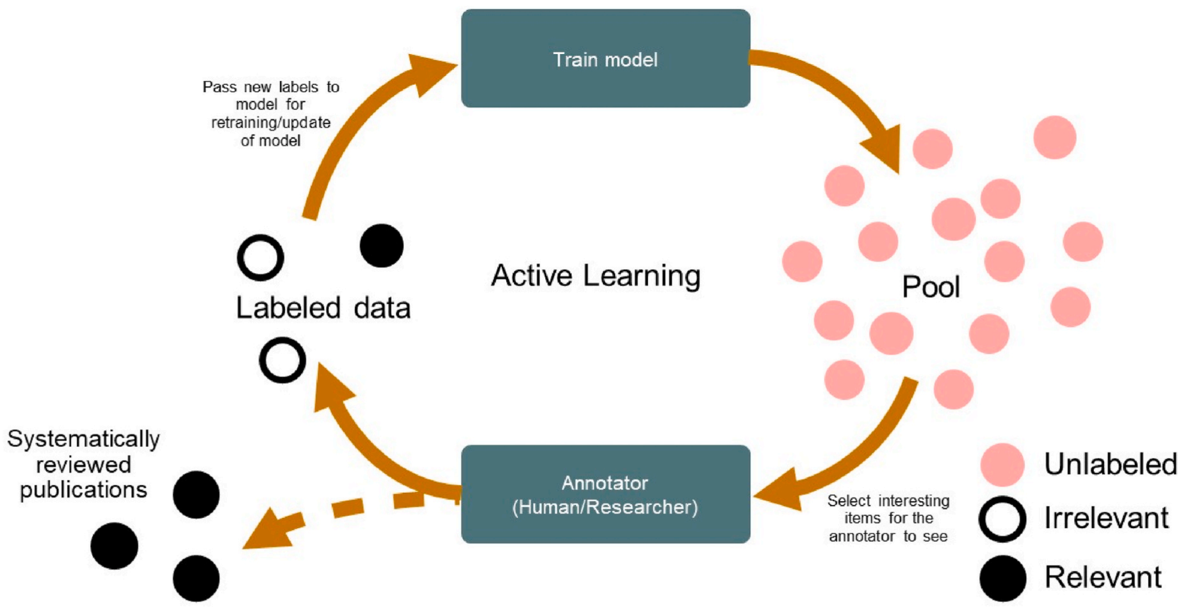

Fig. 2. AI-Aided Pipeline
Note. Figure adopted from https://asreview.readthedocs.io/en/latest/guides/activelearning.html.

corresponding author to provide the full text. We were able to obtain the full texts of 116 studies. Twelve more studies were excluded based on the full-text screening of these 116 studies (nine did not fulfill inclusion criteria [criteria 1 to 7 in Table 2; i.e., in addition to criteria 1 to 6, the full text had to be available]; two were retracted; and one was a duplicate, which we had not previously identified). The flow chart (Fig. 1) also gives an overview of the abstract and full-text screening process until our study's final set of 104 studies was established.

3.3. Coding

The identified 104 studies were then coded by student assistants and the first four authors of this study. The first four authors are experts in the content they have coded. The student assistants were trained intensively (7 h) to use the coding manual. The entire coding manual and coding table can be accessed online in our OSF project: https://doi.org/10.17605/OSF.IO/EMXJQ.

Student assistants coded low-inferential information from the 104 studies, such as publication characteristics (e.g., year of publication, title) and sample characteristics (e.g., age, gender distribution). The experts in the author team coded the variables corresponding to each research question as follows. The two first authors coded information regarding the purpose of the studies (e.g., exploited technological potential of AI used, the target group of individuals of AI, focused construct) to gain insight into how AI applications are used for classroom management processes (RQ1). The third author coded information regarding technical aspects (e.g., features and algorithms used, implementation of technical privacy mechanisms, scalability) to describe what AI applications are used or potentially available for classroom management processes (RQ2). The fourth author coded information about ethical and data security considerations to gain insights into the role of these considerations when using AI applications for classroom management processes (RQ3).

All high-inferential codes (Cooper, 2017; Cooper et al., 2019) were double-coded. For instance, the two first authors coded all studies independently for highly inferential variables regarding how AI applications are used for classroom management processes (RQ1). Interrater reliabilities of $k = 0.62$ (technological potential used) to $k = 0.92$ (target group of individuals of AI applications) emerged, which can be classified as substantial to almost perfect (Landis & Koch, 1977). All discrepancies were discussed between the two first authors, and a consistent solution was developed (consensus validation).

4. Results

4.1. Overview of studies included

Table 3 provides an overview of all included studies. Whereas the included studies were published between 2011 and 2022, 79 % of the articles were published in 2019 or later. Most of the studies were conducted by first authors affiliated with STEM departments (80 %: 84 of 104), whereas only nine publications were conducted by first authors affiliated with an education department (9 %). The remaining studies were conducted by first authors from other departments (12 %: 12 out of 104) or with unknown affiliations (3 %: 3 out of 104). The data set is very diverse in terms of the origins of the authors. Most of the studies are from China (33 %: 34 of 104 studies), the United States (17 %: 18 of 104 studies), and India (11 %: 11 of 104 studies). Researchers from no other country published more than five studies in the data set. 12 % of the studies are from Europe (12 out of 104).

4.2. Educational purpose of AI applications for classroom management

We first analyzed the educational purpose of AI applications for classroom management. That is, we wanted to gain insights into the target group (i.e., students, teachers, or both) of AI applications in the studies, the focus of the AI applications (i.e., cognitive, affective, attendance, behavior), and the exploited technological potentials of AI for classroom management (RQ1). We found that the target group (i.e., the individuals the data algorithms have been trained on and thus about whom the algorithms would provide insights) of AI applications in most studies was students (76 %: 79 of 104 studies). In 16 % of the studies (17 of 104), teachers were the target group; in 8 % (8 of 104), students and teachers were the target groups.

When students were the target group of AI applications, *student behavior* (e.g., standing up, raising hands, head movements) was focused on most frequently (in 30 % of the studies [31 of 104], or respectively in 36 % of the studies [31 of 87] that focused on students). For instance, Luan and Shang (2021) focused on constructing a neural network topology (i.e., Mobilenet-SSD) to analyze classroom actions (e.g., hand

Table 2
Overview of inclusion and exclusion criteria.

No.	Inclusion Criteria	Exclusion Criteria
1	Published online between 2000 and March 2022	Published before 2000
2	English language	Other language than English
3	Classroom setting (one teacher, several students, i.e., interactions involving at least two students)	No classroom settings (e.g., 1:1 situations)
4	Empirical, primary research => Document type: articles, proceedings papers, early access, book chapters, books	Not empirical (Document type: e.g., patents, book reviews, meeting abstracts, editorial materials), not primary research (Document type: e.g., reviews, meta-analysis)
5	Reference to artificial intelligence. AI can be recognized by the following keywords (Note: This list is not to be understood as exhaustive; it is only intended to facilitate the recognition of AI. AI may have been used even if none of these keywords were used): Machine learning Artificial intelligence Classification Classifier Clustering Deep learning Reinforcement learning Supervised learning Unsupervised learning Cross validation Computer vision Natural language processing Random forest Decision trees XGBoost AdaBoost Support vector machine (SVM) Naive Bayes K-nearest neighbor Multi layer perceptron (MLP) Artificial neural network (ANN) Deep neural network (DNN) Convolutional neural network (CNN) Recurrent neural network (RNN) Long-short-term memory (LSTM) Hidden Markov model (HMM) K-means Gaussian mixture model Prediction Accuracy F1 score AUROC (Area under the ROC)	No reference to artificial intelligence. Applications used for assessment and evaluation, profiling and prediction, or intelligent tutoring systems.
6	Purpose: Perception of teaching and learning relevant situations (e.g., attention, cognitive states, emotional states, motivational states, behavior)	Purpose: Applications used for assessments and evaluation or intelligent tutoring systems, learning management systems (LMS), MOOCs.
7	Full text available	Full text not available

raising of students, sitting, writing, sleeping, or playing with a cell phone) using big data, particularly from classroom video recordings. The authors utilized target detection and face recognition algorithms to automate tracking of attendance, engagement, and specific actions such as hand raising, and finally, to improve classroom management (e.g., reducing the need for manual effort) and teaching quality. The algorithms showed good performance, with an average precision of 83.08 %, and improvements in speed and accuracy over traditional methods.

Cognitive aspects (e.g., attention, engagement) and student attendance were each focused on in 24 % of the studies (25 of 104), or respectively in 29 % of the studies (25 of 87) that focused on students. For instance, Goldberg et al. (2021) formulated the idea that teachers, especially novices, could be supported by helping them monitor and understand student engagement more effectively, which would be important to improving teaching quality and learning outcomes, deployed ML algorithms to analyze recorded classroom videos to assess students' engagement through visible behaviors such as gaze direction, head pose, and facial expressions. Specifically, Support Vector Regression was employed to estimate students' engagement based on the features (e.g., head pose) extracted automatically using the OpenFace library. The authors showed that the algorithms perform well, showing strong correlations with manual ratings and learning outcomes. However, manual ratings outperformed automated methods, especially when predicting knowledge test results. Interestingly, it was found that adding synchrony (i.e., considering the behavior of neighboring students) to the model improved the accuracy of the engagement predictions, further highlighting the complexity of classroom interactions. Smith et al. (2019) were driven by the idea that an automated assessment will help educators address attention issues (i.e., determining how many students are attending a lecture and whether they are paying attention) in large classrooms and ultimately improve student success rates. Therefore, the authors employed deep learning methods to monitor classroom occupancy and assess student attention by facial detection and orientation in video data in real-time (i.e., allowing real-time feedback to the educator). Specifically, the authors compared various face detection algorithms (e.g., Haar Cascade, LBP, HOG, ResNet, TinyFace CNN, SSD) to detect student faces in crowded classroom scenes, whereby TinyFace CNN showed the best performance (F1 score of 0.97). Moreover, Convolutional Neural Networks (CNN) models (e.g., AlexNet, VGG19, ResNet101, GoogleNet) were employed to detect whether a student's face is oriented, which helps infer attention levels. VGG19 provided the highest accuracy across different face profiles in real-world classroom scenarios, achieving 80 % accuracy in identifying left, right, and center orientations.

Affective or motivational aspects (e.g., affect states such as boredom, frustration, sleepiness) were focused on least frequently (in 17 % of the studies [18 of 104], or respectively in 21 % of the studies that focused on students [18 of 87]). For instance, Wu (2022) proposed a motion recognition system to detect students' aggressive behaviors early in classroom environments in recorded classroom videos. This might help teachers intervene immediately to prevent potential injuries caused by student conflicts. The author utilized a combination of algorithms and techniques such as background removal (i.e., algorithms to eliminate non-static background elements from video frames), saliency map technology (i.e., highlighting specific regions of interest in an image such as body contours), and motion recognition analysis (i.e., calculating the angles between vectors and tracking the velocity of student movements between consecutive frames). Moreover, the author used an ML technique called confusion matrix and minimum cross-entropy to evaluate and improve the accuracy of aggressive behavior recognition. The proposed algorithm showed excellent performance with high accuracy rates across different student datasets (e.g., US students: 0.98, Taiwan students: 0.98, and Korea students: 0.96).

When teachers were the target group of AI applications, teacher behavior (e.g., hand gestures, facial expressions, body language) was focused on most frequently (in 20 % of the studies [21 of 104], or respectively in 84 % of the studies [21 of 25] that focused on teachers). For instance, Wu et al. (2020) constructed a teacher behavior data set based on online open recorded classroom teaching videos to improve the recognition of teacher behaviors, such as blackboard writing, questioning, displaying, instructing, describing, and non-gestural behavior. By automatically recognizing teacher behaviors in classrooms using multimodal data, the authors aimed to support teaching reflection and reduce information overload when analyzing classroom videos, improve their teaching methods, and support their professional development. The study utilized traditional human action recognition techniques, including RGB Video Analysis (i.e., using dense sampling and feature extraction techniques such as HOG, HOF, and MBH to represent different features such as color or texture), Skeleton Information Extraction (i.e., using the OpenPose algorithm to detect the teacher's

Table 3
Studies included in the systematic review.

Author	Year of publication	Title	Research field of affiliation of first author	Country
Araya & Sossa-Rivera	2021	Automatic detection of gaze and body orientation in elementary school classrooms	Education	Chile
Ashwin & Guddeti	2019	Unobtrusive behavioral analysis of students in classroom environment using non-verbal cues	STEM	India
Ashwin & Guddeti	2020	Automatic detection of students' affective states in classroom environment using hybrid convolutional neural networks	STEM	India
Athanesious et al.	2019	Deep learning based automated attendance system	STEM	India
Bálint	2021	Possibilities for the utilization of an automatized, electronic blockchain-based, students' attendance register, using a universities' modern security cameras	Other	Hungary
Banerjee et al.	2020	Multimodal behavior analysis in computer-enabled laboratories using nonverbal cues	STEM	India
Bassiou et al.	2016	Privacy-preserving speech analytics for automatic assessment of student collaboration	STEM	United States of America
Bhatti et al.	2021	Facial expression recognition of instructor using deep features and extreme learning machine	STEM	Pakistan
Bhavana et al.	2020	Computer vision based classroom attendance management system-with speech output using LBPH algorithm	STEM	India
Bumanis et al.	2020	Children face long-term identification in classroom: Prototype proposal	STEM	Latvia
Calado et al.	2017	A framework to bridge teachers, student's affective state, and improve academic performance	STEM	Portugal
Chen	2021	Evaluation technology of classroom students' learning state based on deep learning	Other	China
Chen et al.	2021	Intelligent teaching evaluation system integrating facial expression and behavior recognition in teaching video	STEM	China
Chiu & Tseng	2021	A Bayesian classification network-based learning status management system in an intelligent classroom	Education	China
Chowdhury et al.	2020	Profiling instructor activities using smartwatch sensors in a classroom	STEM	United States of America
Cui et al.	2021	Machine learning-based student emotion recognition for business English class	na	China
Deniz et al.	2019	Computer vision for attendance and emotion analysis in school settings	STEM	United States of America
Donnelly et al.	2017	Words matter: Automatic detection of teacher questions in live classroom discourse using linguistics, acoustics, and context	STEM	United States of America
Feng et al.	2019	Research on dynamic and static fusion polymorphic gesture recognition algorithm for interactive teaching interface	STEM	China
Fu et al.	2019	Learning behavior analysis in classroom based on deep learning	STEM	China
Giannakos et al.	2020	Fitbit for learning: Towards capturing the learning experience using wearable sensing	STEM	Norway
Gligoric et al.	2015	Smart classroom system for detecting level of interest a lecture creates in a classroom	Other	Serbia and Montenegro
Goldberg et al.	2021	Attentive or not? Toward a machine learning approach to assessing students' visible engagement in classroom instruction	Education	Germany
Guo	2020	Detection of head raising rate of students in classroom based on head posture recognition	na	China
Gupta et al.	2018	CVUCAMS: Computer vision based unobtrusive classroom attendance management system	STEM	India
Gupta et al.	2019	Students' affective content analysis in smart classroom environment using deep learning techniques	STEM	India
Huang & Zhang	2021	Research on learning state based on students' attitude and emotion in class learning	Education	China
Huang et al.	2020	An automatic recognition method for students' classroom behaviors based on image processing	STEM	China
Ismail et al.	2022	Web-based university classroom attendance system based on deep learning face recognition	STEM	Malaysia
James et al.	2018	Inferring the climate in classrooms from audio and video recordings: A machine learning approach	STEM	Singapore
Jensen et al.	2020	Toward automated feedback on teacher discourse to enhance teacher learning	STEM	United States of America
Jisi & Yin	2021	A new feature fusion network for student behavior recognition in education	STEM	China
Kelly et al.	2018	Automatically measuring question authenticity in real-world classrooms	Education	United States of America
Khan et al.	2019	Deep unified model for face recognition based on convolution neural network and edge computing	STEM	Pakistan
Khan et al.	2020	Real time automatic attendance system for face recognition using face API and OpenCV	STEM	Pakistan
Kim et al.	2018	Towards emotionally aware AI smart classroom: Current issues and directions for engineering and education	STEM	United States of America
Klein & Celik	2017	The wits intelligent teaching system: Detecting student engagement during lectures using convolutional neural networks	STEM	South Africa
Ku et al.	2018	Sokrates teaching analytics system (stas): An automatic teaching behavior analysis system for facilitating teacher professional development	STEM + Education	Taiwan
Lasri et al.	2019	Facial emotion recognition of students using convolutional neural network	Other	Marocco
Li	2021	Recognition of psychological characteristics of students' behavior based on improved machine learning	STEM	China
Li et al.	2020	A system for real-time intervention in negative emotional contagion in a smart classroom deployed under edge computing service infrastructure	STEM	China
J. Li et al.	2020	Multimodal learning for classroom activity detection	STEM + Education	China
Li et al.	2021	A convolutional neural network (CNN) based approach for the recognition and evaluation of classroom teaching behavior	STEM	China
Lim et al.	2017	Automated classroom monitoring with connected visioning system	STEM	Malaysia

(continued on next page)

Table 3 (*continued*)

Author	Year of publication	Title	Research field of affiliation of first author	Country
Lin et al.	2021	Student behavior recognition system for the classroom environment based on skeleton pose estimation and person detection	STEM	Taiwan
Liu	2021	Convolutional neural network-assisted strategies for improving teaching quality of college English flipped class	Other	Switzerland
Liu & Liu	2021	A three-dimensional anisotropic diffusion equation-based video recognition model for classroom concentration evaluation in English language teaching	Other	China
Liu et al.	2022	An improved method of identifying learner's behaviors based on deep learning	STEM	China
Luan & Shang	2021	Neural network topology construction and classroom interaction benchmark graph based on big data analysis	na	China
Lukas et al.	2016	Student attendance system in classroom using face recognition technique	STEM	Indonesia
Luo	2021	Artificial intelligence model for real-time monitoring of ideological and political teaching system	STEM	China
Manseras et al.	2018	Millennial Filipino student engagement analyzer using facial feature classification	STEM	Philippines
Martinez-Maldonado et al.	2021	Moodoo the tracker: Spatial classroom analytics for characterising teachers' pedagogical approaches	STEM	Australia
Mehta et al.	2022	Three-dimensional DenseNet self-attention neural network for automatic detection of student's engagement	STEM	India
Menezes et al.	2020	Automatic attendance management system based on Deep One-Shot learning	STEM	Brazil
Mery et al.	2019	Student attendance system in crowded classrooms using a smartphone camera	STEM	Chile
Mindoro et al.	2020	Capturing students' attention through visible behavior: A prediction utilizing YOLOv3 approach	STEM	Philippines
Mitra et al.	2017	Automated student attendance management system using multiple facial images	STEM	Indonesia
Ngoc Anh et al.	2019	A computer-vision based application for student behavior monitoring in classroom	STEM	Vietnam
Nida et al.	2019	Instructor activity recognition through deep spatiotemporal features and feedforward extreme learning machines	STEM	Pakistan
Nishikawa et al.	2015	Estimating positions of students in a classroom from camera images captured by the lecturer's PC	STEM	Japan
Pabba & Kumar	2021	An intelligent system for monitoring students' engagement in large classroom teaching through facial expression recognition	STEM	India
Pei & Shan	2019	A micro-expression recognition algorithm for students in classroom learning based on convolutional neural network	Other	China
Peng et al.	2019	Teaching assistant and class attendance analysis using surveillance camera	STEM	China
Ping et al.	2018	Automatic attendance face recognition for real classroom environments	STEM	China
Prieto et al.	2016	Teaching analytics: Towards automatic extraction of orchestration graphs using wearable sensors	STEM	Switzerland
Prieto et al.	2018	Multimodal teaching analytics: Automated extraction of orchestration graphs from wearable sensor data	STEM	Estonia
Puckdeevongs et al.	2020	Classroom attendance systems based on Bluetooth low energy indoor positioning technology for smart campus	STEM	Thailand
Qiao & Beling	2011	Classroom video assessment and retrieval via multiple instance learning	STEM	United States of America
Rai et al.	2019	An end-to-end real-time face identification and attendance system using convolutional neural networks	STEM	Indonesia
Ramakrishnan et al.	2019	Toward automated classroom observation: Predicting positive and negative climate	STEM	United States of America
Rashmi et al.	2021	Surveillance video analysis for student action recognition and localization inside computer laboratories of a smart campus	STEM	India
Razzaq et al.	2023	DeepClassRooms: A deep learning based digital twin framework for on-campus class rooms	STEM	Pakistan
Renawi et al.	2021	A simplified real-time camera-based attention assessment system for classrooms: pilot study	STEM	United Arab Emirates
Romine et al.	2020	Using machine learning to train a wearable device for measuring students' cognitive load during problem-solving activities based on electrodermal activity, body temperature, and heart rate: Development of a cognitive load tracker for both personal and classroom use	STEM	United States of America
Ross et al.	2013	Using support vector machines to classify student attentiveness for the development of personalized learning systems	STEM	United States of America
Sarkar et al.	2019	Automatic attendance system using deep learning framework	STEM	India
Segal et al.	2017	Keeping the teacher in the loop: Technologies for monitoring group learning in real-time	STEM	Israel
Shao et al.	2018	Multi-object detection based on deep learning in real classrooms	STEM	China
Smith et al.	2019	CNNs and transfer learning for lecture venue occupancy and student attention monitoring	STEM	South Africa
Som et al.	2021	Automated student group collaboration assessment and recommendation system using individual role and behavioral cues	STEM	United States of America
Su et al.	2014	Developing a sensor-based learning concentration detection system	STEM	Taiwan
Su et al.	2021	Learning behaviour recognition based on multi-object image in single viewpoint	STEM	China
Sun et al.	2021	Student class behavior dataset: A video dataset for recognizing, detecting, and captioning students' behaviors in classroom scenes	STEM	China
Suresh et al.	2018	Using deep learning to automatically detect talk moves in teachers'mathematics lessons	STEM	United States of America
Tabassum et al.	2020	Non-intrusive Identification of student attentiveness and finding their correlation with detectable facial emotions	STEM	United States of America
Tang et al.	2019	Design of Intelligent classroom facial recognition based on Deep Learning	Other	China
Uzelac et al.	2015	A comprehensive study of parameters in physical environment that impact students' focus during lecture using Internet of Things	Other	Serbia
van der Haar	2017	Portable computer vision-based cardiac estimation as a teaching aid	STEM	South Africa

(*continued on next page*)

Table 3 (*continued*)

Author	Year of publication	Title	Research field of affiliation of first author	Country
Veliyath et al.	2019	Modeling students' attention in the classroom using eyetrackers	STEM	United States of America
Viswanathan & Vanlehn	2019	Detection of collaboration: Relationship between log and speech-based classification	STEM	United States of America
Viswanathan & VanLehn	2019	Collaboration detection that preserves privacy of students' speech	STEM	United States of America
Wang	2022	A spatio-temporal attention fusion model for students behaviour recognition	Other	China
Wang et al.	2017	Classroom attendance auto-management based on deep learning	STEM	China
Wu	2022	Confusion matrix and minimum cross-entropy metrics based motion recognition system in the classroom	STEM	Taiwan
Wu et al.	2020	The recognition of teacher behavior based on multimodal information fusion	STEM + Education	China
Xie et al.	2021	Abnormal behavior recognition in classroom pose estimation of college students based on spatiotemporal representation learning	STEM	China
Zaletelj	2017	Estimation of students' attention in the classroom from kinect features	STEM	Slovenia
Zaletelj & Košir	2017	Predicting students' attention in the classroom from Kinect facial and body features	STEM	Slovenia
Zhang	2022	Affective cognition of students' autonomous learning in college English teaching based on deep learning	Other	China
Zhang et al.	2022	Intelligent classroom teaching assessment system based on deep learning model face recognition technology	STEM	China
Zhao et al.	2019	The advisable technology of key-point detection and expression recognition for an intelligent class system	Other	China
Zheng et al.	2020	Recognition of teachers' facial expression intensity based on convolutional neural network and attention mechanism	STEM	China
Zylich & Whitehill	2020	Noise-robust key-phrase detectors for automated classroom feedback	STEM	United States of America

skeleton and joint points), and Information Fusion (i.e., combining RGB and skeleton data [multimodal approach] using early fusion [feature-level] and late fusion [decision-level] to improve recognition accuracy). The study reports an accuracy of up to 83.27 % when using early RGB and skeleton data fusion.

Affective or motivational aspects (e.g., sadness) were focused on in 5 % of the studies (5 of 104) or, respectively, in 20 % of the studies (5 of 25) that focused on teachers. Only one study focused on cognitive aspects (i.e., engagement), and no study focused on teachers' attendance.

We found that in 16 % of the studies (15 of 104), technical potentials beyond substitution (i.e., efficiency) were used, meaning that those studies made use of technical potentials that go beyond making coding and analyzing processes more efficient and extend, amplify, or redefine existing approaches. Technical potentials beyond substitution included systems that provide real-time feedback or information (e.g., alerting teachers to critical situations in student interactions; 7 of 15 studies) or recommend specific actions (e.g., feedback options, teacher evaluation; 2 of 15 studies). For instance, Segal et al. (2017) developed a system that aims at supporting teachers in orchestrating multiple groups of students in real-time (i.e., monitor and support collaboratively working on inquiry-based tasks) by providing alerts about critical situations in certain groups; however, the decision to intervene remains with the teacher. The authors used the system SAGLET, which employs supervised learning, specifically a Random Forest model, to classify student utterances such as on-task, off-task, or technical challenges. The Random Forest model performed well, achieving high accuracy for on-task (96.7 %) and technical challenges (75 %), although it had slightly lower accuracy for off-task behaviors (66 %).

Furthermore, studies that deployed optical (e.g., eye-tracking), electroencephalographic, or physical sensors (e.g., gyroscope, accelerometer, Kinect sensors) were also considered to utilize extended technological potentials, as these sensors gather information that cannot be acquired by humans alone (7 of 15 studies). For instance, Giannakos et al. (2020) used wearable devices (i.e., wrist-worn devices [Empatica E4]) to enrich the assessment of students' learning experience by continuously and unobtrusively monitoring students' physiological data during classroom (learning) activities in real-time. The authors used ML algorithms, including Support Vector Machines (SVM) with polynomial and radial kernels, and Random Forests to predict learning experiences based on physiological data (i.e., heart rate, blood volume pulse, electrodermal activity, temperature). The SVM with polynomial kernels showed the best performance, achieving prediction accuracy with an average error rate of 11 % across different learning metrics (e.g., satisfaction, performance, usefulness).

4.3. Technical aspects determining usage of AI applications for classroom management

We next analyzed the features and algorithms used in the studies to gain insights into employed methods (i.e., what is possible from a technical point of view) regarding AI for classroom management (RQ2). Employed features often represent characteristics of students or teachers extracted from various data modalities, such as audio-visual recordings, eye trackers, or other sensors. These features are then input for an ML algorithm to predict the target variable of interest (e.g., student attention). For student outcomes, in Fig. 3, we provide a graphical overview of the approaches (i.e., features used and outcomes focused) of all 104 studies. We can see that most studies using an unimodal approach (i.e., only one feature) used raw image data and focused on students' attendance. Most studies using a multimodal approach (i.e., more than one feature) focused on cognitive student outcomes. Interestingly, an integrated approach (i.e., more than one feature and outcome) was rarely used in general, not at all for student attendance and behavior.

One study may employ a range of feature groups, particularly as combining different sensors, called *multimodality*, can increase the precision and robustness of predictions. This was the case for 28 % of the studies (29 of 104) in our data set. For instance, Chiu and Tseng (2021) developed an AI system to manage and monitor students' learning status in real time. Utilizing multimodal data, including physiological signals (i.e., body temperature, pulse) and data from sensors (e.g., cameras, microphones, temperature sensors), the authors aimed at automatically detecting students' attentiveness and behaviors (e.g., inattention, fatigue based on features such as facial recognition, posture, eye movements) and providing immediate feedback to both teachers and students to enhance focus and engagement, employing a Bayesian classification network. The system demonstrated high accuracy in determining learning statuses, with a statistically significant positive correlation between system predictions and human observer assessments

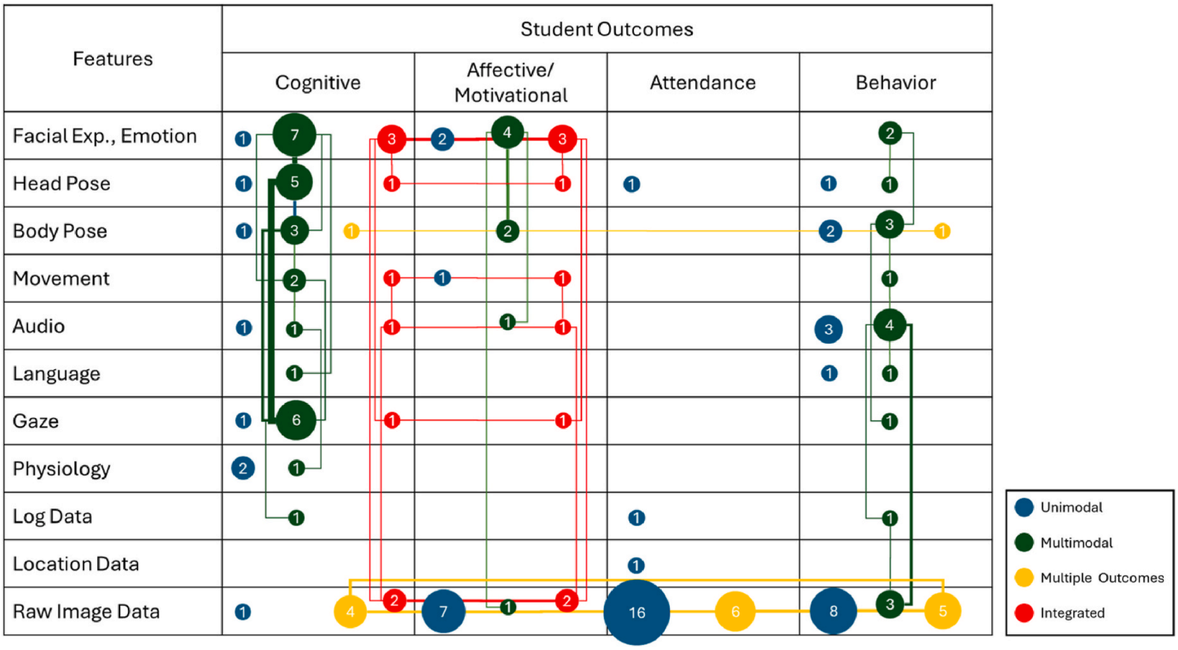

Fig. 3. CMMA (Classroom Management Multimodal Analyses) Grid
Note. Like the illustration by Molenaar et al. (2023), this figure shows different perspectives taken in research on AI used for assessing classroom management (i.e., student outcomes with underlying features). The size of the circles corresponds to the number of associated studies, which are also expressed in numbers. Blue circles represent studies with a unimodal approach (i.e., only one feature used). Green circles and lines represent studies with multimodal approaches (i.e., more than one feature used). Yellow circles and lines represent studies focusing on multiple outcomes (i.e., more than one) but using only one feature. Red circles and lines represent studies with an integrated approach (i.e., more than one feature used and more than one outcome in focus). The size of the circles corresponds to the number of associated studies, which are also expressed in numbers. The thickness of the lines represents the number of associated studies.

(Spearman's rank correlation coefficient of 0.79). Moreover, the system effectively promoted student attention and reduced fatigue during classes.

However, 49 % of the studies (51 of 104) utilized raw video or audio data directly as input for their models, skipping the intermediate feature generation process. For instance, Menezes et al. (2020, pp. 137–142) aimed to streamline attendance assessment, which is typically time-consuming and prone to errors, by using face recognition in real-world classroom settings to verify students' presence with minimal human interaction to automate student attendance. The data consisted of classroom images taken with different devices, where the system detects, aligns, and verifies students' faces. The system employed a deep one-shot learning approach on this data, using a pre-trained FaceNet model based on a deep convolutional neural network (CNN), achieved high accuracy, with an average accuracy of 97.1 % using an iPhone 7, 91.9 % using a Moto G smartphone, and 51.2 % using a low-resolution webcam. The study showed that the performance was highly dependent on the resolution and quality of the camera used for capturing student images.

Table 4
Overview of features used in studies.

Feature Group	Examples	Count	Percentage
Raw image/audio data	Image pixel input	51	49 %
Facial Expressions and Emotions	Facial action units	23	22 %
Acoustic Features	Prosodic, spectral features	16	15 %
Body Pose	Skeleton keypoints	14	13 %
Head Pose	Pitch, yaw and roll	12	12 %
Gaze	Mobile eye tracking	11	11 %
Movement	Motion intensity	8	8 %
Linguistic Features	Presence of question words	8	8 %
(Neuro-) Physiological Features	EDA, EEG, heartrate	4	4 %
Log Data	Tablet user actions	3	3 %
Location	GPS data	3	3 %

An overview of feature groups and frequencies can be found in Table 4. The most leveraged feature group in our sample, employed by 22 % of all studies (23 of 104), was facial expressions and emotions. These were usually extracted from video recordings using an upstream, pre-trained facial expression recognition algorithm. 15 % of the studies (16 of 104) used acoustic features, while 8 % (8 of 104) employed linguistic features, which can, for instance, be extracted from classroom transcripts; four out of these studies combined both feature sets. Also widely used were body pose (13 %: 14 of 104 studies) and head pose (12 %: 12 of 104 studies) features, which were mostly derived from classroom video recordings by employing skeleton key point estimation. Another important feature group found in the data was gaze features, used in 11 % of our sample studies (11 of 104) and obtained mainly by eye tracking or appearance-based gaze estimation techniques. 8 % of the studies (8 of 104) collected movement features, mostly employing accelerometers in wearables such as smartwatches. Less frequently used features were location data in the classroom (3 %: 3 of 104 studies), obtained mainly by GPS trackers, and (neuro-)physiological sensor features (4 %: 4 of 104 studies), such as electrodermal activity (EDA) or EEG measures. 3 % of the studies (3 of 104) conducted in smart-classroom settings used log data from the respective equipment for their AI applications. What is striking is that almost half of the studies in our sample (49 %: 51 of 104) did not employ explicit features but used images from video streams as direct input for deep learning algorithms to predict outcome variables. Often, a preceding step included pre-selection of relevant image areas such as students' faces, i.e., by employing face detection algorithms.

Consequently, it is not surprising that regarding the algorithms employed in AI applications for classroom management, 57 % of the studies (59 of 104) utilized deep learning algorithms such as convolutional neural networks (CNNs), and 12 % (12 of 104) utilized long-short-term memory networks (LSTMs). 24 % of our sample studies (25 of 104) used more traditional, so-called shallow ML algorithms for classification or regression. Such models include random forest classifiers, support vector machines, or k-nearest neighbor classifiers. NLP algorithms were employed in 4 % of the studies (4 of 104). Particularly regarding deep learning algorithms, a clear trend can be observed in the data. In the first three quarters of the analyzed period, 2000 to 2016, about 11 % of the studies (1 of 9) used such algorithms; the proportion in the last quarter, 2017 to 2022, is 67 % (57 of 93). 9 % of our sample studies (9 of 104) did not employ any ML algorithms, implementing rule-based categorizations or summaries of extracted characteristics, for instance, in dashboards.

Regarding the scalability of the proposed ML approaches, only 49 % of the studies (51 of 104) were described as scalable, meaning they are not tailored to a specific scenario. In contrast, 13 % of the studies (14 of 104) mentioned that the employed ML approaches are not scalable.

Another distinction between AI applications is that some are targeted for real-time use in the classroom. In contrast, others are developed for post-hoc analysis or have not yet been tested for real-time performance. We found that 28 % of the studies (29 of 104) implemented online approaches, whereas 46 % (48 of 104) presented offline applications, and the remaining studies did not explicitly state either.

4.4. Ethical considerations regarding AI applications for classroom management

Finally, to gain insight into ethical and data security considerations in the included articles (RQ3), we analyzed all text passages that mentioned respective keywords (e.g., ethics, data) and whether privacy mechanisms were implemented in the AI. Though there are very good reasons to discuss ethics and privacy in this domain, the presence of ethical questions in research on AI in education was limited. Ethical and data security considerations were stated in 22 % of the studies (21 of 104). Studies from India mentioned such concerns in 45 % (5 of 11). They were raised in 25 % of the studies (3 of 12) from Europe, in 22 % (4 of 18) from the United States, and in 13 % (4 of 32) from China. Ethical and data security concerns were not raised in the 27 studies published before 2019 (Fig. 4).

Concerning the implementation of privacy mechanisms, 13 % of the studies (14 of 104) reported such implementations (e.g., not using or restricting video recordings and individual microphones on students or deleting the database after the study). In contrast, in 83 % of the studies (87 of 104), no such privacy-preserving measures were implemented. Among the three countries that contributed the most publications, privacy mechanisms were implemented most frequently by authors from the United States. Whereas about 39 % of the studies from the USA (7 of 18 studies) implemented privacy mechanisms, the proportion in China (1 of 32 studies) and India (1 of 11 studies) was lower, at 3 % and 9 %, respectively.

5. Discussion

In this systematic literature review, we explored AI applications for classroom management and entered a complex terrain characterized by diverse academic foci, technological innovation, and ethical issues. The rise of AI-driven classroom management tools, especially in recent years, is highly dynamic and in line with the recent advances in ML, particularly deep learning and computer vision technologies. Nearly 80 % of the research reviewed in this paper was published between 2019 and March 2022. This massive growth has opened new avenues for understanding and managing classroom dynamics, from tracking student attention and engagement through facial and emotion recognition to analyzing classroom interactions with advanced algorithms.

5.1. Major findings

Our results showed that the focus in the analyzed collection of papers was on student behavior (30 %), student cognition (25 %, e.g., attention), and student emotion (17 %), with the role of AI primarily focused on monitoring and improving student engagement. Only 20 % of the studies focused on teacher behavior. Most studies (84 %) used AI technology merely as a substitution for human coders—without either exploiting it for more advanced purposes (e.g., providing alerts about critical situations, recommending actions, providing feedback) or integrating information from physical sensors (e.g., accelerometer) that human coders would not have at their disposal. Technologies such as those that analyze facial expressions (22 %), acoustic features (15 %), body poses (13 %), head poses (12 %), gaze features (11 %), movement features (8 %), and linguistic features (8 %) are indicative of AI's growing ability to interpret and respond to complex classroom scenarios. In our collection of papers, 28 % used a multimodal approach by relying on multiple classes of features, and 49 % used an approach without explicit features by training deep learning algorithms directly on images from classroom videos. Online applications that allow for real-time use of the data analyses were presented in 29 % of the papers.

Despite these impressive technological advances in recent years, it must be noted that several challenges accompanied the studies we reviewed. The focus has been on the technical aspects of AI-based methods for analyzing and supporting classroom management. This is also evident in the fact that 80 % of the papers we reviewed originated

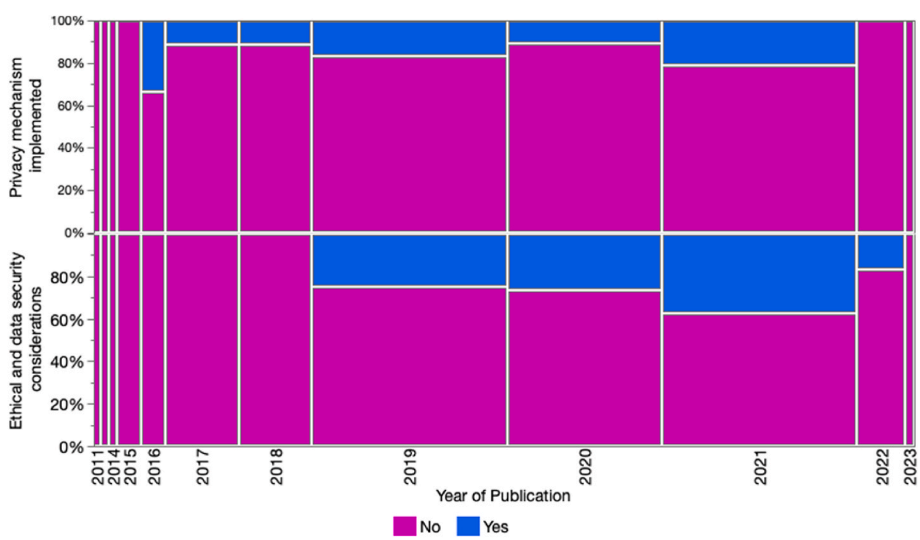

Fig. 4. Ethics, data security, and privacy.

from STEM departments, whereas less than 10 % originated from education departments.

Moreover, whereas ethical considerations of AI applications are equally important as, and should go hand in hand with, technological developments, this is not the case yet. Ethical and data security considerations were stated in only 22 % of the studies, and implementation of privacy mechanisms occurred in only 13 %. In particular, the differences between ethical mindfulness and the implementation of data security measures are striking, especially in different geographical regions. For instance, studies from the United States showed a more pronounced implementation of data protection mechanisms than studies from China and India. On the contrary, ethical and data security considerations were mentioned in 45 % of the studies from India and less often in studies from Europe (25 %), the United States (22 %), and China (13 %). However, in this study, we did not explicitly search for ethics by using the search terms. In other studies, when searching explicitly for ethics, studies from Europe, North America, and Australia tend to be found rather than from China or India (e.g., Hakimi et al., 2021). This geographic variance could reflect the different legal frameworks and cultural attitudes towards data privacy and AI ethics and point to the need for globally consistent ethical standards for the use of AI in education. Therefore, regarding the Wall Street Journal example of AI use in teaching from China shown in the introduction, it can be stated that the associated ethical concerns may be partially justified, but that the reality of the importance of ethical aspects in the studies conducted is rather heterogeneous.

5.2. Ethical recommendations and implications for research and practice

Despite these important limitations, our study highlights some ethical recommendations when applying AI in classrooms (especially regarding classroom management). In general, the increasing use of AI in practice (e.g., in classrooms) requires a balanced approach in which the promise of technological advancement does not overshadow the need for ethical considerations and data security. However, the findings of this literature review illustrate that data protection and ethical aspects are not considered in many AI applications that are (or can be) used in the classroom context (i.e., for classroom management), neither in the AI algorithms themselves nor in the publications of AI applications. These results indicate that many researchers may be unaware of the need for privacy, ethical, and data security measures. Given the sensitive data that (young) students generate in the education sector, we recommend raising awareness among AI developers, whether in companies or research and encouraging them to take data protection and ethical considerations (e.g., data minimization) into account from the beginning of software development. We would like to recommend some good practices for the responsible use of AI in classrooms. First of all, the research on AI in education should improve its standards and the researchers should be mindful of the potential risks such as privacy and data security by obtaining both ethical and data collection permits. Furthermore, it is important that the researchers consider and develop policies about data storage after the studies and their anonymization.

The additional question of best practices for the use of AI in classrooms is a timely one. Strategies for safeguarding the use of AI in schools can come both from the policies which we analyzed, but also from the education itself. Raising awareness of both teachers and parents when it comes to the importance of data privacy can help in requesting responsible AI solutions. When it comes to mitigating the risks of algorithmic biases, awareness and education can help counteract them. In this context, we can consider practicing critical thinking in relation to AI as beneficial. Moreover, addressing the topic of algorithmic fairness has a global component: we want awareness about biased data from a global perspective (Vučković & Sikimić, 2024, pp. 1–18). This can improve the way students, teachers, and parents interpret and use AI outputs and evaluations.

Teaching students how to appropriately use AI becomes one of the important tasks for future education. There is even the fear that the use of genAI decreases the motivation of some students to actually learn the material but overly rely on AI (Walsh, 2025). Technological progress brings many benefits, and we need to educate both minors and adults how to use it in the best way. Awareness of the limitations and appropriate use of AI should be fostered through teaching critical thinking (Hughes, 2024). Thus, one can argue that education itself has the key for resolving the "AI crisis" in education.

Apart from raising awareness among researchers about the sensitivity of AI on children, an incentive system could be created in which journals start requiring ethical permits and data security statements. Cross-national differences would decrease because the policy would come from the scientific editors (i.e., in a top-down approach). We hope that such measures will help to promote the explicit and transparent presentation of data protection measures in publications on AI applications, as this is the only way potential users of AI applications can get a holistic picture of the opportunities and risks and make a balanced decision on their use in education.

Information about AI efficiency and accuracy must be available to make ethical conclusions regarding the ethical principle of beneficence. Beyond research on what is possible technologically, we need research on technology acceptance (do learners and teachers want it) and performance (does it help learning). Therefore, for instance, randomized controlled (field) trials can be used to compare the effectiveness of AI-generated adaptive content with traditional methods (e.g., Meurers et al., 2019). In addition, longitudinal studies could provide valuable insights into the long-term impact of AI on classroom dynamics and learning outcomes.

Regarding the ethical principles of justice and transparency, algorithms used for students' evaluation should not function as black boxes. We recommend that they provide detailed feedback about students' performance to students and teachers so that both parties (as opposed to only the teacher) receive information as starting points for reflection and as a basis for exchange (e.g., about learning behavior).

Conclusions about AI applications for analyzing and supporting classroom management strongly depend on the specific scenarios envisioned for these applications. First, applications for analyzing classroom videos in the context of educational research or teacher training have a completely different flavor than applications used for the online support of teachers in their classroom management. Second, applications that merely inform teachers about potential actions that can be taken are very different from applications that control actions by themselves and are only supervised by teachers (or not; see Molenaar in OECD, 2021). Third, applications running on local machines with federated ML raise different concerns than those running on remote servers. Fourth, it is important to consider students' opportunities to opt out of surveillance or, for specific applications, even consider a requirement for students to opt in actively. These different opportunities to increase human agency in the context of Human-AI-Teaming are very important determinants of how the promises and perils of AI in education will develop in future classrooms.

Finally, not only software developers in companies and research but also policymakers and educators are challenged to develop and adhere to rigorous guidelines that ensure AI is used to improve educational outcomes while protecting the privacy and autonomy of students and teachers. For instance, educators must be trained to use AI in classrooms appropriately (see AI literacy). This includes knowledge and understanding of AI applications, the ability to use and evaluate them, and the ability to recognize and critically weigh ethical aspects (Hornberger et al., 2023; Ng et al., 2021) to critically assess AI applications and make informed decisions about their use in the classroom. Hence, pre-service and in-service teachers must be trained in initial teacher training and professional development programs in the technical aspects of AI applications and in understanding the ethical implications. Here again, AI applications for research and teacher training, on the one hand, and for practical (online) applications, on the other hand, should be distinguished. It is important to distinguish different scenarios, for instance, with or without the opportunity to opt out or the requirement to opt in and with or without local data processing (federated ML) versus server-based analyses that send data to other computers.

5.3. Perspectives for further research

Data validity is highly relevant when using AI in education because measuring classroom management *reliably* and *validly* is a key challenge for researchers. Researchers have drawn on various research methodological approaches in recent decades (for a historical overview, see Brophy, 2006) to assess classroom management strategies' effectiveness (e.g., on student achievement). The consistency of a measure—that is, if the outcomes can be reproduced under the same circumstances—is referred to as reliability. The accuracy of a measure—that is, if the outcomes reflect what the measure is intended to assess—is referred to as validity. Researchers use various methods to assess indicators of classroom management (for an overview see Goe et al., 2008; Lüdtke et al., 2006): *classroom observations* (Bell et al., 2019; White et al., 2024), *teacher self-reports* (e.g., teachers' perception assessed in surveys: e.g., Hamre & Pianta, 2001), or *student self-reports* (e.g., students' perception and feedback assessed in surveys: e.g., Fauth et al., 2014). Whereas classroom observations have the potential to capture an outside perception of classroom management, there are concerns about their reliability (i.e., rater accuracy) and the inferences drawn from individual lessons about teaching in a classroom (see Casabianca et al., 2013; Hill et al., 2012). Teacher self-reports can tap into a teacher's goals, ideas, information, and beliefs about classroom management but may, for instance, be inflated due to social desirability (i.e., the tendency for people to show themselves in a generally positive light; Goe et al., 2008; Moorman & Podsakoff, 1992). Student self-reports can capture individual students' (idiosyncratic) perceptions of classroom management but may be biased by a general impression of a teacher (e.g., teacher popularity). Thus, concerns regarding the reliability and validity of student ratings (e.g., discriminant validity) exist (see Göllner et al., 2018; Wagner et al., 2013). Therefore, AI applications that can capture classroom management not only efficiently but also reliably and validly would open numerous new approaches for research (e.g., analysis of teaching in real-time or of many teaching videos) and practice (e.g., relieving teachers when monitoring the class and thus providing more time to support students). This presupposes that AI systems function reliably and validly, as do human judgments. However, we cannot systematically state how well the AI/ML algorithms work and are generalizable across studies. Therefore, future research should address how the quality of AI/ML algorithms, i.e., their reliability and validity, can be determined *across* studies.

Furthermore, this study focused on privacy as one key general ethical aspect. However, future research should also address the other general ethical aspects of *beneficence*, *justice*, and *transparency* (e.g., Hagendorff, 2022), which are also important to be aware of when applying AI technologies in classrooms. The *principle of beneficence* requires that new teaching solutions based on AI bring real benefits compared to traditional methods, as their use is not justified otherwise. In the research context, beneficence means that AI can only be tested if there are good reasons to assume that its use will have a positive impact. In other words, each research design must consider the comparative benefit of new technologies over the standard ones. For instance, researchers measuring students' attention using AI must provide good evidence for believing that attention trackers are better and more useful than teachers' evaluations or student self-reports. However, such an assessment can only be made if information on the reliability and validity of the AI applications is known. Regarding the *justice* and *transparency* criteria, it is important to ensure that no student is discriminated against during the implementation of AI in classrooms and to guarantee that the evaluation of students is understandable and fair.

Another important aspect is that the EU Artificial Intelligence Act introduces clear legal boundaries for the use of AI in educational contexts. Specifically, the Act prohibits "the placing on the market, the putting into service for this specific purpose, or the use of AI systems to infer emotions of a natural person in the areas of workplace and education institutions," except where such use is intended for medical or safety-related purposes (European Commission, 2021). This prohibition has direct implications for a substantial number of the studies included in this review, many of which relied on emotion recognition technologies for classroom management tasks. At the same time, it is important to note that the regulation includes a research exemption. According to Articles 2 of the Act, AI systems developed and used *exclusively for scientific research purposes* are not subject to the same restrictions—provided they are not placed on the market or made available beyond the research setting. Consequently, academic research on emotion recognition in classrooms may still be permissible under the Act, assuming compliance with other relevant legal and ethical frameworks, such as the General Data Protection Regulation (GDPR), institutional ethical review processes, and informed consent procedures. This distinction is essential for guiding future research and ensuring that innovation continues responsibly within clearly defined legal boundaries.

Finally, future AI applications for classroom management—particularly those involving real-time, multimodal data—could benefit from parallel computing frameworks to handle increasing data complexity and processing demands. Recent advances in parallel deep learning architectures (e.g., Khan, Akram, & Usman, 2020; Noor et al., 2025) demonstrate the potential of distributed systems for efficiently modeling high-dimensional data, offering promising directions for scalable and responsive educational AI systems.

Against this backdrop, we initiated the Emerging Field Group (EFG) 13 of the European Association for Research on Learning and Instruction (EARLI) *Automated Assessment of Teaching Effectiveness Using Multimodal Data* (https://www.earli.org/efg-13) to address these questions and future research directions through a global and interdisciplinary collaboration.

5.4. Limitations

An important limitation of this study is the uncertainty of the generalizability of the results. Several factors might affect the generalizability of our findings. First, it is well known that positive or successful outcomes of AI applications (e.g., AI applications that work regarding predictive reliability and validity) are more easily and quickly published (i.e., publication bias) or that authors tend to report selected findings (i.e., reporting bias). As we cannot rule out such biases in the included studies of this review, there may be concerns about the generalizability of our findings, particularly regarding their applicability in different educational settings and cultural contexts—a challenge typically known in ML and computer vision applications. Furthermore, the overrepresentation of certain countries in the data set (i.e., USA and China) also limits the broader applicability of our conclusions, as unique challenges and opportunities in different educational landscapes may be overlooked. Second, most of the research we report on in this paper goes back to 2019 to 2021, presenting only a snapshot of a very dynamic development with rapid accelerations. Our findings, particularly on the neglect of ethical considerations, might hopefully be obsolete once the general awareness of the promises and perils of AI in education rises, particularly in the more technically oriented research areas, due to societal discourse.

Another limitation of this study is the focus on traditional AI and the possibly insufficient consideration of current trends in genAI. The landscape of AI in education is becoming even more complex with the emergence of generative AI technologies, which are analytical and capable of creating content. This opens new possibilities for classroom management and learning methods. However, these applications of genAI in classroom management remain less explored, presenting a frontier rich with potential yet to be fully understood and harnessed.

Finally, as this review did not involve the development or testing of AI models, we were unable to assess the computational complexity or conduct benchmarking analyses based on specific datasets, which limits our ability to evaluate the technical efficiency and generalizability of the included systems.

6. Conclusion

The findings reveal that current AI applications for classroom management that are reported to perform well (e.g., reliable predictions) are predominantly utilized to monitor and enhance student engagement. On the one hand, technological advances in AI, such as facial and emotion recognition and processing of multimodal data, offer promising avenues for managing classroom dynamics to ensure effective teaching. On the other hand, regarding the threat that current AI applications in education could also be used to control students, the review highlights a notable gap in ethical mindfulness and data security. Therefore, we highlight the need to balance technological advancement with ethical considerations, data security, and privacy. Responsible and ethical use of AI applications is paramount to safeguard students' and teachers' privacy and autonomy and ensure equitable and transparent educational practices. Against this backdrop, we call for increased awareness among AI developers (e.g., the necessity of incorporating data protection measures from the initial stages of software development) and researchers (e.g., the importance of transparent reporting on AI applications). The review also underscores the need for further research on AI's efficiency, acceptance, and impact on learning outcomes to ensure its beneficence in educational settings.

CRediT authorship contribution statement

Tim Fütterer: Writing – original draft, Project administration, Methodology, Formal analysis, Conceptualization. **Patricia Goldberg:** Writing – original draft, Project administration, Conceptualization. **Babette Bühler:** Writing – original draft, Visualization, Formal analysis. **Vlasta Sikimić:** Writing – original draft, Formal analysis. **Ulrich Trautwein:** Writing – review & editing. **Peter Gerjets:** Writing – review & editing. **Kathleen Stürmer:** Writing – review & editing. **Enkelejda Kasneci:** Writing – review & editing.

Declaration

We have no known conflicts of interest to disclose. The authors are responsible for the content of this publication.

Declaration of competing interest

We have no known conflicts of interest to disclose. Data and Supplementary Material of the study can be accessed at our Open Science Framework (OSF) project: https://doi.org/10.17605/OSF.IO/EMXJQ. The authors are responsible for the contents of this publication.

We used Grammarly (1.98.2.0) and ChatGPT-4o to perform language editing in the manuscript.

Acknowledgments

This research was supported by the Tübingen Postdoctoral Academy for Research on Education (PACE) of the Hector Research Institute of Education Sciences and Psychology, Tübingen, funded by the Baden-Württemberg Ministry of Science, Research, and the Arts.

References

References marked with an asterisk were included in the systematic review.

* Araya, R., & Sossa-Rivera, J. (2021). Automatic detection of gaze and body orientation in elementary school classrooms. *Frontiers in Robotics and AI, 8*. https://doi.org/10.3389/frobt.2021.729832.

* Ashwin, T. S., & Guddeti, R. M. R. (2019). Unobtrusive behavioral analysis of students in classroom environment using non-verbal cues. *IEEE Access, 7*, 150693–150709. https://doi.org/10.1109/ACCESS.2019.2947519.

* Ashwin, T. S., & Guddeti, R. M. R. (2020). Automatic detection of students' affective states in classroom environment using hybrid convolutional neural networks. *Education and Information Technologies, 25*(2), 1387–1415. https://doi.org/10.1007/s10639-019-10004-6.

* Athanesious, J. J., Vanitha, Adithya, S., Bhardwaj, C. A., Lamba, J. S., & Vaidehi, A. V. (2019). Deep learning based automated attendance system. *2nd International Conference on Recent Trends in Advanced Computing ICRTAC -DISRUP - TIV INNOVATION, 165*, 307–313. https://doi.org/10.1016/j.procs.2020.01.045, 2019 November 11-12, 2019.

Baker, T., & Smith, L. (2019). *Educ-AI-tion rebooted? Exploring the future of artificial intelligence in schools and colleges*. nesta. https://media.nesta.org.uk/documents/Future_of_AI_and_education_v5_WEB.pdf.

* Bálint, K. (2021). Possibilities for the utilization of an automatized, electronic blockchain-based, students' attendance register, using a universities' modern security cameras. *Acta Polytechnica Hungarica, 18*(2), 127–142. https://doi.org/10.12700/APH.18.2.2021.2.7.

* Banerjee, S., Ashwin, T. S., & Guddeti, R. M. R. (2020). Multimodal behavior analysis in computer-enabled laboratories using nonverbal cues. *Signal, Image and Video Processing, 14*(8), 1617–1624. https://doi.org/10.1007/s11760-020-01705-4.

* Bassiou, N., Tsiartas, A., Smith, J., Bratt, H., Richey, C., Shriberg, E., D'Angelo, C., & Alozie, N. (2016). *Privacy-preserving speech analytics for automatic assessment of student collaboration* (pp. 888–892).

Bell, C. A., Dobbelaer, M. J., Klette, K., & Visscher, A. (2019). Qualities of classroom observation systems. *School Effectiveness and School Improvement, 30*(1), 3–29. https://doi.org/10.1080/09243453.2018.1539014

* Bhatti, Y. K., Jamil, A., Nida, N., Yousaf, M. H., Viriri, S., & Velastin, S. A. (2021). Facial expression recognition of instructor using deep features and extreme learning machine. *Computational Intelligence and Neuroscience, 2021*, Article 5570870. https://doi.org/10.1155/2021/5570870.

* Bhavana, D., Kumar, K. K., Kaushik, N., Lokesh, G., Harish, P., Mounisha, E., & Tej, D. R. (2020). Computer vision based classroom attendance management system-with speech output using LBPH algorithm. *International Journal of Speech Technology, 23*(4), 779–787. https://doi.org/10.1007/s10772-020-09739-2.

Brakel, M., & Uuk, R. (n.d.). The artificial intelligence act. The AI Act. https://artificialintelligenceact.eu/.

Brophy, J. E. (1983). Classroom organization and management. *The Elementary School Journal, 83*(4), 264–285. https://doi.org/10.1086/461318

Brophy, J. E. (2006). History of research on classroom management. In C. M. Evertson, & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice, and contemporary issues* (pp. 17–43). Lawrence Erlbaum Associates Publishers.

* Bumanis, N., Vitols, G., Arhipova, I., & Meirane, I. (2020). Children face long-term identification in classroom: Prototype proposal. *Proceedings of the 22nd International Conference on Enterprise Information Systems (ICEIS 2020), 2*, 287–293. https://doi.org/10.5220/0009414002870293.

* Calado, J., Luís-Ferreira, F., Sarraipa, J., & Jardim-Goncalves, R. (2017). A framework to bridge teachers, student's affective state, and improve academic performance. In *ASME 2017 international mechanical engineering congress and exposition* (Vol. 2). https://doi.org/10.1115/IMECE2017-72000.

Campos, D. G., Fütterer, T., Gfrörer, T., Lavelle-Hill, R., Murayama, K., König, L., … Scherer, R. (2024). Screening smarter, not harder: A comparative analysis of machine learning screening algorithms and heuristic stopping criteria for systematic reviews in educational research. *Educational Psychology Review, 36*(1). https://doi.org/10.1007/s10648-024-09862-5.

Casabianca, J. M., McCaffrey, D. F., Gitomer, D. H., Bell, C. A., Hamre, B. K., & Pianta, R. C. (2013). Effect of observation mode on measures of secondary mathematics teaching. *Educational and Psychological Measurement, 73*(5), 757–783. https://doi.org/10.1177/0013164413486987

* Chen, L. (2021). Evaluation technology of classroom students' learning state based on deep learning. *Computational Intelligence and Neuroscience, 2021*, Article 6999347. https://doi.org/10.1155/2021/6999347.

* Chen, Z., Liang, M., Yu, W., Huang, Y., & Wang, X. (2021). Intelligent teaching evaluation system integrating facial expression and behavior recognition in teaching video. *2021 IEEE international conference on big data and smart computing (BigComp)* (pp. 52–59). https://doi.org/10.1109/BigComp51126.2021.00019.

* Chiu, C.-K., & Tseng, J. C. R. (2021). A Bayesian classification network-based learning status management system in an intelligent classroom. *Educational Technology & Society, 24*(3), 256–267. JSTOR.

* Chowdhury, Z. U., De, P., & Allen, A. A. (2020). Profiling instructor activities using smartwatch sensors in a classroom. *Proceedings of the 2020 ACM southeast conference* (pp. 135–140). https://doi.org/10.1145/3374135.3385290.

Cooper, H. M. (2017). *Research synthesis and meta-analysis: A step-by-step approach* (5th ed.). Sage.

Cooper, H. M., Hedges, L. V., & Valentine, J. C. (Eds.). (2019). *Handbook of research synthesis and meta-analysis* (3rd ed.). Russell Sage Foundation.

* Cui, Y., Wang, S., & Zhao, R. (2021). Machine learning-based student emotion recognition for business English class. *International Journal of Emerging Technologies in Learning (iJET), 16*(12), 94–107. https://doi.org/10.3991/ijet.v16i12.23313.

* Deniz, S., Lee, D., Kurian, G., Altamirano, L., Yee, D., Ferra, M., Hament, B., Zhan, J., Gewali, L., & Oh, P. (2019). Computer vision for attendance and emotion analysis in school settings. *2019 IEEE 9th annual computing and communication workshop and conference (CCWC)* (pp. 134–139). https://doi.org/10.1109/CCWC.2019.8666488.

Deutscher Ethikrat. (2023). Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz [Stellungnahme] https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf.
Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM, 55*(10), 78–87. https://doi.org/10.1145/2347736.2347755
* Donnelly, P. J., Blanchard, N., Olney, A. M., Kelly, S., Nystrand, M., & D'Mello, S. K. (2017). Words matter: Automatic detection of teacher questions in live classroom discourse using linguistics, acoustics, and context. *Proceedings of the seventh international learning analytics & knowledge conference* (pp. 218–227). https://doi.org/10.1145/3027385.3027417.
Emmer, E. T., & Evertson, C. M. (1981). Synthesis of research on classroom management. *Educational Leadership, 38*, 342–347.
European Commission. (2021). *Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (2021)*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206.
Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction, 29*, 1–9. https://doi.org/10.1016/j.learninstruc.2013.07.001
* Feng, Z., Xu, T., Yang, X., Tian, J., Yi, J., & Zhao, K. (2019). Research on dynamic and static fusion polymorphic gesture recognition algorithm for interactive teaching interface. In F. Sun, H. Liu, & D. Hu (Eds.), *Cognitive systems and signal processing* (pp. 104–115). Singapore: Springer. https://doi.org/10.1007/978-981-13-7986-4_10.
Foster, J. K., Korban, M., Youngs, P., Watson, G. S., & Acton, S. T. (2024). Automatic classification of activities in classroom videos. *Computers and Education: Artificial Intelligence, 6*, 100207. https://doi.org/10.1016/j.caeai.2024.100207.
* Fu, R., Wu, T., Luo, Z., Duan, F., Qiao, X., & Guo, P. (2019). Learning behavior analysis in classroom based on deep learning. *2019 tenth international conference on intelligent control and information processing (ICICIP)* (pp. 206–212). https://doi.org/10.1109/ICICIP47338.2019.9012177.
* Giannakos, M. N., Sharma, K., Papavlasopoulou, S., Pappas, I. O., & Kostakos, V. (2020). Fitbit for learning: Towards capturing the learning experience using wearable sensing. *International Journal of Human-Computer Studies, 136*, Article 102384. https://doi.org/10.1016/j.ijhcs.2019.102384.
* Gligoric, N., Uzelac, A., Krco, S., Kovacevic, I., & Nikodijevic, A. (2015). Smart classroom system for detecting level of interest a lecture creates in a classroom. *Journal of Ambient Intelligence and Smart Environments, 7*(2), 271–284. https://doi.org/10.3233/AIS-150303.
Goe, L., Bell, C. A., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. National Comprehensive Center for Teacher Quality.
* Goldberg, P., Sümer, Ö., Stürmer, K., Wagner, W., Göllner, R., Gerjets, P., Kasneci, E., & Trautwein, U. (2021). Attentive or not? Toward a machine learning approach to assessing students' visible engagement in classroom instruction. *Educational Psychology Review, 33*(1), 27–49. https://doi.org/10.1007/s10648-019-09514-z.
Göllner, R., Wagner, W., Eccles, J. S., & Trautwein, U. (2018). Students' idiosyncratic perceptions of teaching quality in mathematics: A result of rater tendency alone or an expression of dyadic effects between students and teachers? *Journal of Educational Psychology, 110*(5), 709–725. https://doi.org/10.1037/edu0000236
Grossman, P. (Ed.). (2018). *Teaching core practices in teacher education*. Harvard Education Press.
* Guo, Q. (2020). Detection of head raising rate of students in classroom based on head posture recognition. *Traitement du Signal, 37*(5). https://doi.org/10.18280/ts.370515.
* Gupta, S. K., Ashwin, T. S., & Guddeti, R. M. R. (2018). CVUCAMS: Computer vision based unobtrusive classroom attendance management system. *2018 IEEE 18th international conference on advanced learning technologies (ICALT)* (pp. 101–102). https://doi.org/10.1109/ICALT.2018.00131.
* Gupta, S. K., Ashwin, T. S., & Guddeti, R. M. R. (2019). Students' affective content analysis in smart classroom environment using deep learning techniques. *Multimedia Tools and Applications, 78*(18), 25321–25348. https://doi.org/10.1007/s11042-019-7651-z.
Hagendorff, T. (2022). A virtue-based framework to support putting AI ethics into practice. *Philosophy & Technology, 35*(3), 55. https://doi.org/10.1007/s13347-022-00553-z
Hakimi, L., Eynon, R., & Murphy, V. A. (2021). The ethics of using digital trace data in education: A thematic review of the research landscape. *Review of Educational Research, 91*(5), 671–717. https://doi.org/10.3102/00346543211020116
Hamre, B. K., & Pianta, R. C. (2001). Early teacher-child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development, 72*(2), 625–638. https://doi.org/10.1111/1467-8624.00301
Hardy, I., Decristan, J., & Klieme, E. (2019). Adaptive teaching in research on learning and instruction. *Journal for Educational Research Online, 11*, 169–191.
Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
Hill, H. C., Charalambous, C. Y., & Kraft, M. A. (2012). When rater reliability is not enough: Teacher observation systems and a case for the generalizability study. *Educational Researcher, 41*(2), 56–64. https://doi.org/10.3102/0013189X12437203
Holzinger, A. (2016). Interactive machine learning for health informatics: When do we need the human-in-the-loop? *Brain Informatics, 3*(2), 119–131. https://doi.org/10.1007/s40708-016-0042-6
Hornberger, M., Bewersdorff, A., & Nerdel, C. (2023). What do university students know about artificial intelligence? Development and validation of an AI literacy test. *Computers and Education: Artificial Intelligence, 5*, Article 100165. https://doi.org/10.1016/j.caeai.2023.100165
Huang, W., Li, N., Qiu, Z., Jiang, N., Wu, B., & Liu, B. (2020). An automatic recognition method for students' classroom behaviors based on image processing. *Traitement du Signal, 37*(3), 503–509. https://doi.org/10.18280/ts.370318
Hou, R., Fütterer, T., Bühler, B., Bozkir, E., Gerjets, P., Trautwein, U., & Kasneci, E. (2024). In A. M. Olney, I.-A. Chounta, Z. Liu, O. C. Santos, & I. I. Bittencourt (Eds.), *Artificial Intelligence in Education, 14829* pp. 60–74). Switzerland: Springer Nature. https://doi.org/10.1007/978-3-031-64302-6_5.
* Huang, D., & Zhang, W. (2021). Research on learning state based on students' attitude and emotion in class learning. *Scientific Programming, 2021*, Article 9944176. https://doi.org/10.1155/2021/9944176.
Hughes, C. (2024). *Critical thinking and generative artificial intelligence*. UNESCO International Bureau of Education. Retrieved May 23, 2025, from https://www.ibe.unesco.org/en/articles/critical-thinking-and-generative-artificial-intelligence.
Hutson, M. (2023). Rules to keep AI in check: Nations carve different paths for tech regulation. A guide to how China, the EU and the US are reining in artificial intelligence. *Nature, 620*, 260–263. https://doi.org/10.1038/d41586-023-02491-y
* Ismail, N. A., Chai, C. W., Samma, H., Salam, M. S., Hasan, L., Wahab, A., Haliza, N., Mohamed, F., Leng, W. Y., & Rohani, M. F. (2022). Web-based university classroom attendance system based on deep learning face recognition. *KSII Transactions on Internet & Information Systems, 16*(2). https://doi.org/10.3837/tiis.2022.02.008.
* James, A., Kashyap, M., Chua, Y. H. V., Maszczyk, T., Núñez, A. M., Bull, R., & Dauwels, J. (2018). Inferring the climate in classrooms from audio and video recordings: A machine learning approach. *2018 IEEE international conference on teaching, assessment, and learning for engineering (TALE)* (pp. 983–988). https://doi.org/10.1109/TALE.2018.8615327.
* Jensen, E., Dale, M., Donnelly, P. J., Stone, C., Kelly, S., Godley, A., & D'Mello, S. K. (2020). Toward automated feedback on teacher discourse to enhance teacher learning. *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1–13). https://doi.org/10.1145/3313831.3376418.
* Jisi, A., & Yin, S. (2021). A new feature fusion network for student behavior recognition in education. *Journal of Applied Science and Engineering, 24*(2), 133–140. https://doi.org/10.6180/jase.202104_24(2).0002.
* Kelly, S., Olney, A. M., Donnelly, P., Nystrand, M., & D'Mello, S. K. (2018). Automatically measuring question authenticity in real-world classrooms. *Educational Researcher, 47*(7), 451–464. https://doi.org/10.3102/0013189X18785613.
* Khan, S., Akram, A., & Usman, N. (2020). Real time automatic attendance system for face recognition using face API and OpenCV. *Wireless Personal Communications, 113* (1), 469–480. https://doi.org/10.1007/s11277-020-07224-2.
* Khan, M. Z., Harous, S., Hassan, S. U., Khan, M. U. G., Iqbal, R., & Mumtaz, S. (2019). Deep unified model for face recognition based on convolution neural network and edge computing. *IEEE Access, 7*, 72622–72633. https://doi.org/10.1109/ACCESS.2019.2918275.
Khan, S., Khan, M., Iqbal, N., Li, M., & Khan, D. M. (2020). Spark-based parallel neep neural network model for classification of large scale RNAs into piRNAs and non-piRNAs. *IEEE Access, 8*, 136978–136991. https://doi.org/10.1109/ACCESS.2020.3011508
Khan, S., Noor, S., Javed, T., Naseem, A., Aslam, F., AlQahtani, S. A., & Ahmad, N. (2025). XGBoost-enhanced ensemble model using discriminative hybrid features for the prediction of sumoylation sites. *BioData Mining, 18*(1), 12. https://doi.org/10.1186/s13040-024-00415-8
Khan, S., Uddin, I., Noor, S., AlQahtani, S. A., & Ahmad, N. (2025). N6-methyladenine identification using deep learning and discriminative feature integration. *BMC Medical Genomics, 18*(1), 58. https://doi.org/10.1186/s12920-025-02131-6
* Kim, Y., Soyata, T., & Behnagh, R. F. (2018). Towards emotionally aware AI smart classroom: Current issues and directions for engineering and education. *IEEE Access, 6*, 5308–5331. https://doi.org/10.1109/ACCESS.2018.2791861.
* Klein, R., & Celik, T. (2017). The wits intelligent teaching system: Detecting student engagement during lectures using convolutional neural networks. *2017 IEEE international conference on image processing (ICIP)* (pp. 2856–2860). https://doi.org/10.1109/ICIP.2017.8296804.
Klieme, E., Pauli, C., & Reusser, K. (2009). The pythagoras study: Investigating effects of teaching and learning in swiss and german mathematics classrooms. In T. Janík, & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (pp. 137–160). Waxmann.
Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research, 86*(3), 643–680. https://doi.org/10.3102/0034654315626799
Kounin, J. S. (1970). *Discipline and group management in classrooms*. Holt, Rinehart & Winston.
* Ku, O., Liang, J.-K., Chang, S.-B., & Wu, M. (2018). *Sokrates teaching analytics system (stas): An automatic teaching behavior analysis system for facilitating teacher professional development* (pp. 696–705).
Kunter, M., Baumert, J., & Köller, O. (2007). Effective classroom management and the development of subject-related interest. *Learning and Instruction, 17*(5), 494–509. https://doi.org/10.1016/j.learninstruc.2007.09.002
Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159. https://doi.org/10.2307/2529310
* Lasri, I., Solh, A. R., & Belkacemi, M. E. (2019). Facial emotion recognition of students using convolutional neural network. *2019 third international conference on intelligent computing in data sciences (ICDS)* (pp. 1–6). https://doi.org/10.1109/ICDS47004.2019.8942386.
Leonelli, S. (2015). What counts as scientific data? A relational framework. *Philosophy of Science, 82*(5), 810–821. https://doi.org/10.1086/684083

Lewis, T. J., & Sugai, G. (1999). Effective behavior support: A systems approach to proactive schoolwide management. *Focus on Exceptional Children, 31*(6). https://doi.org/10.17161/foec.v31i6.6767

* Li, M. (2021). Recognition of psychological characteristics of students' behavior based on improved machine learning. *Journal of Sensors, 2021*, Article 8135942. https://doi.org/10.1155/2021/8135942.

* Li, H., Kang, Y., Ding, W., Yang, S., Yang, S., Huang, G. Y., & Liu, Z. (2020). Multimodal learning for classroom activity detection. *Icassp 2020 - 2020 IEEE international conference on acoustics, speech and signal processing* (pp. 9234–9238). ICASSP). https://doi.org/10.1109/ICASSP40776.2020.9054407.

* Li, G., Liu, F., Wang, Y., Guo, Y., Xiao, L., & Zhu, L. (2021). A convolutional neural network (CNN) based approach for the recognition and evaluation of classroom teaching behavior. *Scientific Programming.* , Article 6336773. https://doi.org/10.1155/2021/6336773, 2021.

* Li, J., Shi, D., Tumnark, P., & Xu, H. (2020). A system for real-time intervention in negative emotional contagion in a smart classroom deployed under edge computing service infrastructure. *Peer-to-Peer Networking and Applications, 13*(5), 1706–1719. https://doi.org/10.1007/s12083-019-00863-8.

* Lim, J. H., Teh, E. Y., Geh, M. H., & Lim, C. H. (2017). Automated classroom monitoring with connected visioning system. *2017 Asia-Pacific signal and information processing association annual summit and conference (APSIPA ASC)* (pp. 386–393). https://doi.org/10.1109/APSIPA.2017.8282063.

* Lin, F.-C., Ngo, H.-H., Dow, C.-R., Lam, K.-H., & Le, H. L. (2021). Student behavior recognition system for the classroom environment based on skeleton pose estimation and person detection. *Sensors, 21*(16). https://doi.org/10.3390/s21165314.

* Liu, T. (2021). Convolutional neural network-assisted strategies for improving teaching quality of college English flipped class. *Wireless Communications and Mobile Computing.* , Article 1929077. https://doi.org/10.1155/2021/1929077, 2021.

* Liu, Y., & Liu, J. (2021). A three-dimensional anisotropic diffusion equation-based video recognition model for classroom concentration evaluation in English language teaching. *Advances in Mathematical Physics, 2021*, Article 2209526. https://doi.org/10.1155/2021/2209526.

* Liu, S., Zhang, J., & Su, W. (2022). An improved method of identifying learner's behaviors based on deep learning. *The Journal of Supercomputing, 78*(10), 12861–12872. https://doi.org/10.1007/s11227-022-04402-w.

* Luan, C., & Shang, P. (2021). Neural network topology construction and classroom interaction benchmark graph based on big data analysis. *Wireless Communications and Mobile Computing, 2021*, Article 2334443. https://doi.org/10.1155/2021/2334443.

Lüdtke, O., Trautwein, U., Kunter, M., & Baumert, J. (2006). Reliability and agreement of student ratings of the classroom environment: A reanalysis of TIMSS data. *Learning Environments Research, 9*(3), 215–230. https://doi.org/10.1007/s10984-006-9014-8

* Lukas, S., Mitra, A. R., Desanti, R. I., & Krisnadi, D. (2016). Student attendance system in classroom using face recognition technique. *2016 international conference on information and communication technology convergence (ICTC)* (pp. 1032–1035). https://doi.org/10.1109/ICTC.2016.7763360.

* Luo, Y. (2021). Artificial intelligence model for real-time monitoring of ideological and political teaching system. *Journal of Intelligent and Fuzzy Systems, 40*(2), 3585–3594. https://doi.org/10.3233/JIFS-189394.

* Manseras, R., Eugenio, F., & Palaoag, T. (2018). Millennial Filipino student engagement analyzer using facial feature classification. *IOP Conference Series: Materials Science and Engineering, 325*(1), Article 012006. https://doi.org/10.1088/1757-899X/325/1/012006.

* Martinez-Maldonado, R., Echeverria, V., Mangaroska, K., Shibani, A., Fernandez-Nieto, G., Schulte, J., & Buckingham Shum, S. (2022). Moodoo the tracker: Spatial classroom analytics for characterising teachers' pedagogical approaches. *International Journal of Artificial Intelligence in Education, 32*(4), 1025–1051. https://doi.org/10.1007/s40593-021-00276-w.

* Mehta, N. K., Prasad, S. S., Saurav, S., Saini, R., & Singh, S. (2022). Three-dimensional DenseNet self-attention neural network for automatic detection of student's engagement. *Applied Intelligence, 52*(12), 13803–13823. https://doi.org/10.1007/s10489-022-03200-4.

* Menezes, A. G., Sá, J. M. D.d. C., Llapa, E., & Estombelo-Montesco, C. A. (2020). Automatic attendance management system based on deep one-shot learning. *2020 international conference on systems, signals and image processing (IWSSIP)*. https://doi.org/10.1109/IWSSIP48289.2020.9145230.

* Mery, D., Mackenney, I., & Villalobos, E. (2019). Student attendance system in crowded classrooms using a smartphone camera. *2019 IEEE winter conference on applications of computer vision (WACV)* (pp. 857–866). https://doi.org/10.1109/WACV.2019.00096.

Meurers, D., De Kuthy, K., Nuxoll, F., Rudzewitz, B., & Ziai, R. (2019). Scaling up intervention studies to investigate real-life foreign language learning in school. *Annual Review of Applied Linguistics, 39*, 161–188. https://doi.org/10.1017/S0267190519000126

* Mindoro, J. N., Pilueta, N. U., Austria, Y. D., Lacatan, L. L., & Dellosa, R. M. (2020). Capturing students' attention through visible behavior: A prediction utilizing YOLOv3 approach. *2020 11th IEEE control and system graduate research colloquium (ICSGRC)* (pp. 328–333). https://doi.org/10.1109/ICSGRC49013.2020.9232659.

* Mitra, A. R., Lukas, S., Desanti, R. I., & Krisnadi, D. (2017). Automated student attendance management system using multiple facial images. *2nd international conference on computational modeling, simulation and applied mathematics (CMSAM 2017)*. https://doi.org/10.12783/dtcse/cmsam2017/16422.

Molenaar, I., Mooij, S. D., Azevedo, R., Bannert, M., Järvelä, S., & Gasevic, D. (2023). Measuring self-regulated learning and the role of AI: Five years of research using multimodal multichannel data. *Computers in Human Behavior, 139*, 107540. https://doi.org/10.1016/j.chb.2022.107540.

Moorman, R. H., & Podsakoff, P. M. (1992). A meta-analytic review and empirical test of the potential confounding effects of social desirability response sets in organizational behaviour research. *Journal of Occupational and Organizational Psychology, 65*(2), 131–149. https://doi.org/10.1111/j.2044-8325.1992.tb00490.x

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence, 2*, Article 100041. https://doi.org/10.1016/j.caeai.2021.100041

* Ngoc Anh, B., Tung Son, N., Truong Lam, P., Phuong Chi, L., Huu Tuan, N., Cong Dat, N., Huu Trung, N., Umar Aftab, M., & Van Dinh, T. (2019). A computer-vision based application for student behavior monitoring in classroom. *Applied Sciences, 9*(22). https://doi.org/10.3390/app9224729.

* Nida, N., Yousaf, M. H., Irtaza, A., & Velastin, S. A. (2019). Instructor activity recognition through deep spatiotemporal features and feedforward extreme learning machines. *Mathematical Problems in Engineering.* , Article 2474865. https://doi.org/10.1155/2019/2474865, 2019.

* Nishikawa, J., Kakusho, K., Iiyama, M., Nishiguchi, S., & Murakami, M. (2015). Estimating positions of students in a classroom from camera images captured by the lecturer's PC. In N. Streitz, & P. Markopoulos (Eds.), *Lecture notes in computer science: 9189. Distributed, ambient, and pervasive interactions. DAPI 2015* (pp. 518–526). Springer International Publishing. https://doi.org/10.1007/978-3-319-20804-6_47.

Noor, S., Awan, H. H., Hashmi, A. S., Saeed, A., Khan, S., & AlQahtani, S. A. (2025). Optimizing performance of parallel computing platforms for large-scale genome data analysis. *Computing, 107*(3), 86. https://doi.org/10.1007/s00607-025-01441-y

OECD. (2021). *OECD digital education outlook 2021: Pushing the frontiers with artificial intelligence, blockchain and robots*. OECD. https://doi.org/10.1787/589b283f-en

* Pabba, C., & Kumar, P. (2022). An intelligent system for monitoring students' engagement in large classroom teaching through facial expression recognition. *Expert Systems, 39*(1), Article e12839. https://doi.org/10.1111/exsy.12839.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 1–9. https://doi.org/10.1136/bmj.n71

Panchal, K., & Mohammad, B. S. (2020). Artificial intelligence used in school's of China. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3560832

Patall, E. A., Yates, N., Lee, J., Chen, M., Bhat, B. H., Lee, K., Beretvas, S. N., Lin, S., Man Yang, S., Jacobson, N. G., Harris, E., & Hanson, D. J. (2023). A meta-analysis of teachers' provision of structure in the classroom and students' academic competence beliefs, engagement, and achievement. *Educational Psychologist*, 1–29. https://doi.org/10.1080/00461520.2023.2274104

* Pei, J., & Shan, P. (2019). A micro-expression recognition algorithm for students in classroom learning based on convolutional neural network. *Traitement du Signal, 36* (6), 557–563. https://doi.org/10.18280/ts.360611.

* Peng, X., Fang, Z., & Gao, Y. (2019). Teaching assistant and class attendance analysis using surveillance camera. In G. Zhai, J. Zhou, P. An, & X. Yang (Eds.). *Digital TV and multimedia communication. IFTC 2018. Communications in computer and information science, 1009* pp. 413–422). Springer. https://doi.org/10.1007/978-981-13-8138-6_35.

* Ping, C., Da-Peng, H., & Zu-Ying, L. (2018). Automatic attendance face recognition for real classroom environments. *Proceedings of the 2018 2nd international conference on big data and internet of things* (pp. 65–70). https://doi.org/10.1145/3289430.3289436.

Posit team. (2022). RStudio: Integrated development for R (2022.02.1) [Computer software]. Posit Software, PBC http://www.posit.co/.

Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of three basic dimensions. *ZDM, 50*(3), 407–426. https://doi.org/10.1007/s11858-018-0918-4

* Prieto, L. P., Sharma, K., Dillenbourg, P., & Jesús, M. (2016). Teaching analytics: Towards automatic extraction of orchestration graphs using wearable sensors. *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 148–157). https://doi.org/10.1145/2883851.2883927.

* Prieto, L. P., Sharma, K., Kidzinski, Ł., Rodríguez-Triana, M. J., & Dillenbourg, P. (2018). Multimodal teaching analytics: Automated extraction of orchestration graphs from wearable sensor data. *Journal of Computer Assisted Learning, 34*(2), 193–203. https://doi.org/10.1111/jcal.12232.

* Puckdeevongs, A., Tripathi, N. K., Witayangkurn, A., & Saengudomlert, P. (2020). Classroom attendance systems based on bluetooth low energy indoor positioning technology for smart campus. *Information, 11*(6). https://doi.org/10.3390/info11060329.

* Qiao, Q., & Beling, P. A. (2011). Classroom video assessment and retrieval via multiple instance learning. In G. Biswas, S. Bull, J. Kay, & A. Mitrovic (Eds.), *Lecture notes in computer science: 6738. Artificial intelligence in education* (pp. 272–279). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-21869-9_36. *AIED 2011*.

R Core Team. (2022). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing [Computer software] https://www.R-project.org.

* Rai, A., Karnani, R., Chudasama, V., & Upla, K. (2019). An end-to-end real-time face identification and attendance system using convolutional neural networks. *2019 IEEE 16th India council international conference (INDICON)* (pp. 1–4). https://doi.org/10.1109/INDICON47234.2019.9029001.

* Ramakrishnan, A., Ottmar, E., LoCasale-Crouch, J., & Whitehill, J. (2019). Toward automated classroom observation: Predicting positive and negative climate. *2019 14th IEEE international conference on automatic face & gesture recognition (FG 2019)* (pp. 1–8). https://doi.org/10.1109/FG.2019.8756529.

* Rashmi, M., Ashwin, T. S., & Guddeti, R. M. R. (2021). Surveillance video analysis for student action recognition and localization inside computer laboratories of a smart campus. *Multimedia Tools and Applications, 80*(2), 2907–2929. https://doi.org/10.1007/s11042-020-09741-5.
* Razzaq, S., Shah, B., Iqbal, F., Ilyas, M., Maqbool, F., & Rocha, A. (2023). DeepClassRooms: A deep learning based digital twin framework for on-campus class rooms. *Neural Computing & Applications, 35*(11), 8017–8026. https://doi.org/10.1007/s00521-021-06754-5.
Regulation of the European parliament and of the council laying Down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206, (2021).
Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-Networks. https://doi.org/10.48550/ARXIV.1908.10084.
* Renawi, A., Alnajjar, F., Parambil, M., Trabelsi, Z., Gochoo, M., Khalid, S., & Mubin, O. (2021). A simplified real-time camera-based attention assessment system for classrooms: Pilot study. *Education and Information Technologies, 27*(4), 4753–4770. https://doi.org/10.1007/s10639-021-10808-5.
Robert Bosch, S. (2023). *Das Deutsche Schulbarometer: Aktuelle Herausforderungen aus Sicht der Lehrkräfte. Ergebnisse einer Befragung von Lehrkräften allgemein- und berufsbildender Schulen*. Robert Bosch Stiftung. https://deutsches-schulportal.de/content/uploads/2023/09/Schulbarometer_Lehrkraefte_2023_FACTSHEET.pdf.
* Romine, W. L., Schroeder, N. L., Graft, J., Yang, F., Sadeghi, R., Zabihimayvan, M., Kadariya, D., & Banerjee, T. (2020). Using machine learning to train a wearable device for measuring students' cognitive load during problem-solving activities based on electrodermal activity, body temperature, and heart rate: Development of a cognitive load tracker for both personal and classroom use. *Sensors, 20*(17). https://doi.org/10.3390/s20174833.
* Ross, M., Graves, C. A., Campbell, J. W., & Kim, J. H. (2013). Using support vector machines to classify student attentiveness for the development of personalized learning systems. In *2013 12th international conference on machine learning and applications* (Vol. 1, pp. 325–328). https://doi.org/10.1109/ICMLA.2013.66.
* Sarkar, P. R., Mishra, D., & Subhramanyam, G. R. K. S. (2019). Automatic attendance system using deep learning framework. In M. Tanveer, & R. B. Pachori (Eds.), *Machine intelligence and signal analysis* (pp. 335–346). Singapore: Springer. https://doi.org/10.1007/978-981-13-0923-6_29.
* Segal, A., Hindi, S., Prusak, N., Swidan, O., Livni, A., Palatnic, A., Schwarz, B., & Gal, Y. (K.). (2017). Keeping the teacher in the loop: Technologies for monitoring group learning in real-time. In E. André, R. Baker, X. Hu, M. M. T. Rodrigo, & B. du Boulay (Eds.), *Artificial intelligence in education* (pp. 64–76). Springer International Publishing. https://doi.org/10.1007/978-3-319-61425-0_6.
Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research, 77*(4), 454–499. https://doi.org/10.3102/0034654307310317
* Shao, B., Jiang, F., & Shen, R. (2018). Multi-object detection based on deep learning in real classrooms. In X. Geng, & B.-H. Kang (Eds.), *Pricai 2018: Trends in artificial intelligence* (pp. 352–359). Springer International Publishing. https://doi.org/10.1007/978-3-319-97310-4_40.
* Smith, A. J., Van Wyk, B. J., & Du, S. (2019). CNNs and transfer learning for lecture venue occupancy and student attention monitoring. In G. Bebis, R. Boyle, B. Parvin, D. Koracin, D. Ushizima, S. Chai, S. Sueda, X. Lin, A. Lu, D. Thalmann, C. Wang, & P. Xu (Eds.), *Advances in visual computing* (pp. 383–393). Springer International Publishing. https://doi.org/10.1007/978-3-030-33723-0_31.
* Som, A., Kim, S., Lopez-Prado, B., Dhamija, S., Alozie, N., & Tamrakar, A. (2021). Automated student group collaboration assessment and recommendation system using individual role and behavioral cues. *Frontiers of Computer Science, 3*. https://doi.org/10.3389/fcomp.2021.728801.
* Su, Y.-N., Hsu, C.-C., Chen, H.-C., Huang, K.-K., & Huang, Y.-M. (2014). Developing a sensor-based learning concentration detection system. *Engineering Computations, 31*(2), 216–230. https://doi.org/10.1108/EC-01-2013-0010.
* Su, K., Li, X., Zhou, C., & Chen, X. (2021). Learning behaviour recognition based on multi-object image in single viewpoint. *Personal and Ubiquitous Computing, 25*(6), 1081–1090. https://doi.org/10.1007/s00779-019-01286-1.
* Sun, B., Wu, Y., Zhao, K., He, J., Yu, L., Yan, H., & Luo, A. (2021). Student class behavior dataset: A video dataset for recognizing, detecting, and captioning students' behaviors in classroom scenes. *Neural Computing & Applications, 33*(14), 8335–8354. https://doi.org/10.1007/s00521-020-05587-y.
* Suresh, A., Sumner, T., Huang, I., Jacobs, J., Foland, B., & Ward, W. (2018). Using deep learning to automatically detect talk moves in teachers'mathematics lessons. *2018 IEEE international conference on big data (big data)* (pp. 5445–5447). https://doi.org/10.1109/BigData.2018.8621901.
* Tabassum, T., Allen, A. A., & De, P. (2020). Non-intrusive identification of student attentiveness and finding their correlation with detectable facial emotions. *Proceedings of the 2020 ACM southeast conference* (pp. 127–134). https://doi.org/10.1145/3374135.3385263.
* Tang, J., Zhou, X., & Zheng, J. (2019). *Design of intelligent classroom facial recognition based on deep learning, 1168*, Article 022043. https://doi.org/10.1088/1742-6596/1168/2/022043, 2.
* Uzelac, A., Gligoric, N., & Krco, S. (2015). A comprehensive study of parameters in physical environment that impact students' focus during lecture using internet of things. *Computers in Human Behavior, 53*, 427–434. https://doi.org/10.1016/j.chb.2015.07.023.
Van de Schoot, R., & De Bruin, J. (2020). Researcher-in-the-loop for systematic reviewing of text databases. https://doi.org/10.5281/ZENODO.4013207.
van de Schoot, R., de Bruin, J., Schram, R., Zahedi, P., de Boer, J., Weijdema, F., Kramer, B., Huijts, M., Hoogerwerf, M., Ferdinands, G., Harkema, A., Willemsen, J., Ma, Y., Fang, Q., Hindriks, S., Tummers, L., & Oberski, D. L. (2021). An open source machine learning framework for efficient and transparent systematic reviews. *Nature Machine Intelligence, 3*(2), 125–133. https://doi.org/10.1038/s42256-020-00287-7
* van der Haar, D. T. (2017). Portable computer vision-based cardiac estimation as a teaching aid. *Proceedings of the 3rd international conference on communication and information processing* (pp. 439–443). https://doi.org/10.1145/3162957.3163006.
* Veliyath, N., De, P., Allen, A. A., Hodges, C. B., & Mitra, A. (2019). Modeling students' attention in the classroom using eyetrackers. *Proceedings of the 2019 ACM southeast conference* (pp. 2–9). https://doi.org/10.1145/3299815.3314424.
Vincent-Lancrin, S., & van der Vlies, R. (2020). Trustworthy artificial intelligence (AI) in education. In *OECD education working papers, 218*. https://doi.org/10.1787/a6c90fa9-en
* Viswanathan, S. A., & VanLehn, K. (2019). Collaboration detection that preserves privacy of students' speech. In S. Isotani, E. Millán, A. Ogan, P. Hastings, B. McLaren, & R. Luckin (Eds.), *Artificial intelligence in education* (pp. 507–517). Springer International Publishing. https://doi.org/10.1007/978-3-030-23204-7_42.
Viswanathan, S. A., & Vanlehn, K. (2019). Detection of collaboration: Relationship between log and speech-based classification. In S. Isotani, E. Millán, A. Ogan, P. Hastings, B. McLaren, & R. Luckin (Eds.), *Artificial intelligence in education* (pp. 327–331). Springer International Publishing. https://doi.org/10.1007/978-3-030-23207-8_60.
Vučković, A., & Sikimić, V. (2024). *Global justice and the use of AI in education: Ethical and epistemic aspects*. AI & SOCIETY.
Wagner, W., Göllner, R., Helmke, A., Trautwein, U., & Lüdtke, O. (2013). Construct validity of student perceptions of instructional quality is high, but not perfect: Dimensionality and generalizability of domain-independent assessments. *Learning and Instruction, 28*, 1–11. https://doi.org/10.1016/j.learninstruc.2013.03.003
Walsh, J. D. (2025). *Everyone is cheating their way through college*. New York Times Magazine. https://nymag.com/intelligencer/article/openai-chatgpt-ai-cheating-education-college-students-school.html.
* Wang, X. (2022). A spatio-temporal attention fusion model for students behaviour recognition. *EAI Endorsed Transactions on Scalable Information Systems, 9*(34), e1. https://doi.org/10.4108/eai.3-9-2021.170905.
* Wang, D., Fu, R., & Luo, Z. (2017). Classroom attendance auto-management based on deep learning. *Proceedings of the 2017 2nd international conference on education, sports, arts and management engineering (ICESAME 2017)* (pp. 1523–1528). https://doi.org/10.2991/icesame-17.2017.327.
Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research, 63*(3), 249–294. https://doi.org/10.3102/00346543063003249
* Wu, M.-T. (2022). Confusion matrix and minimum cross-entropy metrics based motion recognition system in the classroom. *Scientific Reports, 12*(1), 3095. https://doi.org/10.1038/s41598-022-07137-z.
White, M., Edelsbrunner, P. A., & Thurn, C. (2024). The conceptualization implies the statistical model: Implications for measuring domains of teaching quality. *Assessment in Education: Principles, Policy & Practice, 31*(3–4), 254–278. https://doi.org/10.1080/0969594X.2024.2368252.
* Wu, D., Chen, J., Deng, W., Wei, Y., Luo, H., & Wei, Y. (2020). The recognition of teacher behavior based on multimodal information fusion. *Mathematical Problems in Engineering, 2020*, Article 8269683. https://doi.org/10.1155/2020/8269683.
* Xie, Y., Zhang, S., & Liu, Y. (2021). Abnormal behavior recognition in classroom pose estimation of college students based on spatiotemporal representation learning. *Traitement du Signal, 38*(1). https://doi.org/10.18280/ts.380109.
* Zaletelj, J. (2017). Estimation of students' attention in the classroom from kinect features. *Proceedings of the 10th international symposium on image and signal processing and analysis* (pp. 220–224). https://doi.org/10.1109/ISPA.2017.8073599.
* Zaletelj, J., & Košir, A. (2017). Predicting students' attention in the classroom from kinect facial and body features. *EURASIP Journal on Image and Video Processing, 2017*(1), 80. https://doi.org/10.1186/s13640-017-0228-8.
Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0
* Zhang, D. (2022). Affective cognition of students' autonomous learning in college English teaching based on deep learning. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.808434.
* Zhang, X., Zhang, X., & Jasni Bin Dolah, T. D. (2022). Intelligent classroom teaching assessment system based on deep learning model face recognition technology. *Scientific Programming, 2022*, Article 1851409. https://doi.org/10.1155/2022/1851409.
* Zhao, Y., Yan, H., & Wang, Z. (2019). The advisable technology of key-point detection and expression recognition for an intelligent class system. *Journal of Physics: Conference Series, 1187*(5), Article 052011. https://doi.org/10.1088/1742-6596/1187/5/052011.
* Zheng, K., Yang, D., Liu, J., & Cui, J. (2020). Recognition of teachers' facial expression intensity based on convolutional neural network and attention mechanism. *IEEE Access, 8*, 226437–226444. https://doi.org/10.1109/ACCESS.2020.3046225.
* Zylich, B., & Whitehill, J. (2020). Noise-robust key-phrase detectors for automated classroom feedback. *Icassp 2020 - 2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 9215–9219). https://doi.org/10.1109/ICASSP40776.2020.9053173.