

Article

Not peer-reviewed version

Learning Analytics in the Era of Large Language Models

Elisabetta Mazzullo , [Okan Bulut](#) ^{*} , [Tarid Wongvorachan](#) , Bin Tan

Posted Date: 4 August 2023

doi: 10.20944/preprints202308.0366.v1

Keywords: learning analytics; large language model; artificial intelligence; process data



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Learning Analytics in the Era of Large Language Models

Elisabetta Mazzullo ¹, Okan Bulut ^{2*}, Tarid Wongvorachan ¹ and Bin Tan ¹

¹ Measurement, Evaluation, and Data Science, University of Alberta, Edmonton, AB T6G 2G5, Canada; mazzullo@ualberta.ca; wongvora@ualberta.ca; btan4@ualberta.ca

² Centre for Research in Applied Measurement and Evaluation, University of Alberta, Edmonton, AB T6G 2G5, Canada; bulut@ualberta.ca

* Author to whom correspondence should be addressed.

Abstract: Although learning analytics (LA) holds great potential to improve teaching and learning, LA research and practice are currently riddled with limitations that impact every stage of the LA life cycle. The present paper offers an overview of these challenges before proposing strategies to overcome them and exploring how the recent innovations brought forth by language models can improve LA research and practice. In particular, we encourage the empowerment of teachers during LA development, as this would strengthen the theoretical foundation of LA solutions and increase their interpretability and usability. Furthermore, we provide examples of how process data can be used to understand learning processes and generate more interpretable LA insights. Furthermore, we explore how LLMs could come into play in LA to generate interpretable insights, timely and actionable feedback, increase personalization, and support teachers' tasks more broadly.

Keywords: learning analytics; large language model; artificial intelligence; process data

1. Introduction

Rapid technological advancements are bringing about significant transformations in every aspect of the education system. The advent of digital learning environments (DLE) has made large volumes of novel data available. As students interact in the DLE, "digital traces" about learning, performance, and engagement are recorded [1]. To exploit these new forms of information and make use of computational analysis techniques, Learning Analytics (LA) has emerged as a new research field at the intersection between students' learning, data analytics, and human-centered design. LA is defined as the "measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environment in which it occurs" [2, p.4]. For example, many efforts in LA have been devoted to information visualization or predicting students' academic performance. The key utility of LA as listed by the Society for Learning Analytics Research [SoLAR] [1] include: 1) promoting the development of learning skills and strategies; 2) offering personalized and timely feedback; 3) increasing student awareness by supporting self-reflection; and 4) generating empirical evidence on the success of pedagogical innovations. With an increasing number of studies being published each year [3], the research field was soon recognized as holding great potential to improve learning by supporting students and educators.

Although researchers have generated a vast literature on LA over the past decade, there are still pitfalls in the development of LA systems relative to their theoretical grounding and design choices [4–6], challenges in their implementation [7], and issues in the evaluation of their effectiveness [8]. Jivet *et al.* [9] recognize these as key "moments" in the life-cycle of LA, which should always be informed by learning theories to produce valuable tools. More recently, scholars have shifted their attention to raising awareness of the importance of making teachers an integral part of the LA design process and improving the usability of LA systems based on learning theories [10]. Moreover, with the release of AI systems that can complete a variety of tasks, from memorizing basic concepts to generating narratives and ideas using human-like language, technology is revolutionizing the way we

think about learning and opening up new standards for teaching practices [11]. Based on the discussed shortcomings in LA practices, this paper offers an overview of the current challenges and limitations and proposes directions for its future development. In particular, we encourage teacher empowerment while developing LA systems and using LA to aid teaching practices. To this end, we reflect on how process data and large language models (LLMs) could be harnessed to address some of the pitfalls of LA and to support instructional tasks.

This paper begins by introducing various types of LA and their applications. Then, we present the challenges that modern LA practices face. Figure 1 provides a visual overview of these limitations, situating them within the life-cycle of LA, together with their proposed solutions. Insufficient grounding in learning sciences and poor design choices during the development of LA systems exacerbate issues in the interpretability of LA insights, which add to further challenges in their implementation related to prediction and actionability of feedback. Lastly, the evaluation of LA solutions brings forward issues related to their generalizability and scarce evidence of their effectiveness. From the issues presented, we put forward our recommendations based on the existing literature to involve teachers as LA designers for interpretable pedagogy-based LA systems. We also recommend using process data and natural language processing (NLP) to enhance the interpretability of LA. After that, we discuss how natural language models and their larger variants, like ChatGPT, can increase LA personalization and support teaching practices. We conclude the paper by discussing how the posited recommendations can enhance LA practice as a whole.

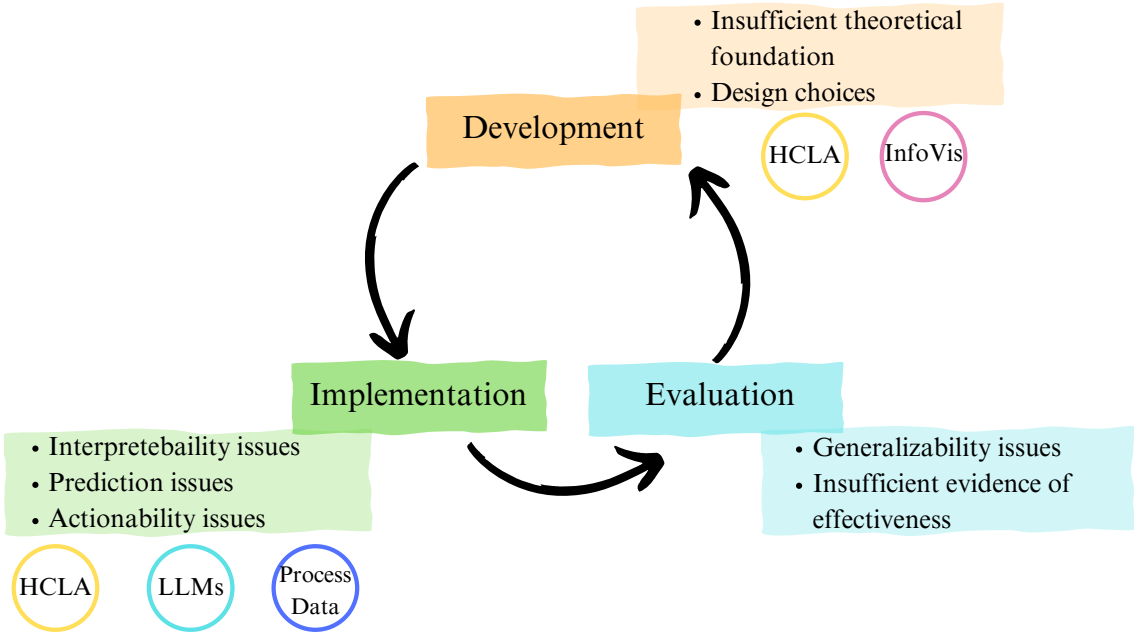


Figure 1. Limitations in the life-cycle of LA and proposed solutions. Note: Circles represent the proposed solutions. HCLA stands for Human-Centered Learning Analytics. InfoVis stands for Information Visualization. LLM stands for Large Language Models.

2. LA: Limitations and Ongoing Challenges

This section briefly presents the different scopes of existing LA systems and illustrates the weaknesses of current research and practices in LA.

2.1. Descriptive, Predictive, and Prescriptive LA

Insights from LA systems are often communicated to stakeholders through LA dashboards (LADs), “a single display that aggregates different indicators about learners, learning processes and learning contexts into one or multiple visualizations” [12, p.37]. LADs can display multiple types

of information. Descriptive analytics show trends and relationships among learning indicators (e.g., grades and engagement compared to peers). Descriptive dashboards typically provide performance visualizations and outcome feedback [13]. Applying modern computational techniques to educational data, researchers are attempting to answer not only how students are doing but also why they performed as they did, how they are expected to perform, or what they should do going forward. Predictive LA systems use machine learning algorithms that use current and past data patterns to forecast future outcomes. Academic outcomes most typically predicted by LADs include grades in upcoming assignments and final exams, risk of non-submission, failing the course, or dropping out. As further explored below, predictive analytics come with their own set of technical limitations and ethical challenges. More recently, the interest is shifting towards the development of prescriptive dashboards offering process-oriented feedback: actionable recommendations pointing students to what they should be doing next to reach their learning goals [13–16]. Examples of similar systems can be found in the “call to action” emails employed by Iraj *et al.* [17], or in a LAD providing students with content recommendations and skill-building activities [8].

2.2. Insufficient Grounding in Learning Sciences

Many scholars now have criticized existing LA systems for their insufficient grounding in the learning sciences and call for a better balance between theory and data [4,5]. Most studies, especially at the beginning of LA investigations, assumed a data-driven approach without adopting specific learning theories to guide their analysis. This way, behavioral patterns can be identified, but their interpretation and understanding remain problematic [5]. The same definition of LA identifies measurement and analytics not as the goal itself but as “means to the end” [18], which is the understanding and optimization of learning and educational environments. This implies that the data is meaningful only to the extent it supports interpretation and guides action.

Of the 49 articles included in their literature review, Algayres and Triantafyllou [19] found that only 28 presented a theoretical framework, primarily referring to theories of self-regulated learning. Similarly, a scoping review on theory-driven LA articles published between 2016 and 2020 yielded 37 studies that revealed that the most commonly used theories include self-regulated learning and social constructivism [5]. The authors invite researchers to explore behavioral and cognitive theories, going beyond observable behavioral log data and investigating information processing strategies (e.g., problem-solving, memory). Moreover, they highlight how learning theories should be used to interpret LA data and promote pedagogical advancement by validating learning designs. DLEs make an astounding amount of data available to researchers and educators. Still, without a theory, they are left astray in interpreting them and deciding which variables are valuable and should be selected for their models [20]. Furthermore, sometimes aggregate measures derived from simple indicators from the log data are more informative for learning, as they better represent learning behaviors studied by educational theories [21]. Therefore, it is essential to understand what these new measures generated in the DLE mean and to remember that engagement does not necessarily equate to learning [22].

2.3. Interpretability Challenges

Interpretability of insights derived from LA is not only related to the theory that is (or is not) underlying the data but also to the choices being made related to communication and design. LADs are located at the intersection between educational data science and information visualization. Recently, scholars have been reminding LADs researchers and developers that these instruments should not merely display data and ask students and teachers to assume the role of data scientists; instead, their main goal should be communicating the most essential information [23]. Research shows that learners’ ability to interpret data may be limited, and, to best support cognition, design choices should be founded on principles of cognitive psychology and information visualization [6]. For example, coherent displays and colors can reduce visual clutter and direct attention to the most essential elements for correctly interpreting the data. The usability and interpretability of LA tools are

critical. In general, educators feel that LA fosters their professional development [24]; however, even if LA tools are perceived as valuable, teachers sometimes struggle to translate data into actions [25]. According to the Technology-Acceptance Model [26], the intention to use technology is influenced by its perceived usefulness and ease of use of the instrument. Therefore, even though teachers recognize the potential benefits of LA, they might avoid using dashboards if they don't feel comfortable navigating or interpreting them.

2.4. Prediction Issues

While they provide richer information than purely descriptive LA systems, predictive models come with their own set of limitations and ethical challenges, such as the risk of stereotyping and biased forecasts. Evidence on teacher use of predictive LA tools is mixed; some studies find that teachers who make more intense and consistent use of this tool can better identify students who need additional support [27], while others do not corroborate these findings [28]. Furthermore, predictions are often generated by black-box models, lacking transparency, interpretability, and explicability [29]. These characteristics favor actionability [30], and in their absence utility of the system and users' trust are reduced [31]. Researchers highlight the need to improve prediction accuracy, together with its validity and generalizability [32], and advise that predictions need to be followed by appropriate actions and effective interventions to influence final results [33].

In the second edition of the Handbook of Learning Analytics, SoLAR provides directions for the use of measurement to transition from predictive models to explanatory models. The goal of LA is optimization, which goes a step further than prediction. At the same time, explanation is neither necessary nor sufficient for optimization; there needs to be a causal mechanism on which students and teachers base their decisions if these actions are expected to produce specific desired outcomes [34].

2.5. Beyond Prediction: Actionability Issue for Automatically Generated Feedback

Researchers are now advocating for the development of LADs that inform students not only about how they have performed so far but also how they can do better [15,29]. As it is widely recognized, feedback supports learning and academic achievement [35]. Earlier studies on feedback adopted an information paradigm, focusing on the type of information provided to learners, its precision, and level of cognitive complexity [17,35]. More recently, the focus has shifted to feedback as a dialogic process and its actionability: students (and teachers) are not passive recipients of information. However, it is crucial to develop their abilities to understand feedback and take action [36]. For feedback to be effective, learners need to understand the information, evaluate their own work, manage their emotions related to the feedback, and take appropriate actions [36,37].

Another important characteristic of good feedback is timeliness. Research shows that the effectiveness of feedback is greater when it is received quickly [38]. LA tools can offer instructors and learners constant access to automatic-generated feedback and real-time monitoring of their performance. Iraj *et al.* [17] found early engagement with feedback to be positively associated with student outcomes when instructors used an LA tool to monitor students' progress and send personalized weekly emails that provided learners with feedback on their activity and highlighted the actions required next in their learning through "call to action" links to task materials. Prescriptive information is appreciated by students [13,14] and seems to support student motivation [39]. However, emerging prescriptive dashboards often rely on human intervention or employ automated algorithms based on hard-core heuristics and thresholds, so some researchers call for developing more sophisticated systems [15].

Moreover, the effectiveness of feedback is influenced by student characteristics [35,40], and it is enhanced when feedback is personalized [17]. LA tools can offer instructors and learners constant access to automatic-generated feedback and real-time monitoring of their performance, and they offer new opportunities to provide individualized feedback to students. Technology-mediated feedback systems have been found to increase students' engagement, satisfaction, and outcomes [41,42]. By

favoring personalization and timeliness of feedback, together with the display of adequate and actionable information, student-facing LADs could help reduce the “feedback gap” [17], the difference between the potential and actual use of feedback [41,43]. However, an extensive literature review from Matcha *et al.* [4] suggests that existing dashboards, with their scarce grounding in theory, are unlikely to follow literature recommendations for best feedback.

2.6. Generalizability Issue

The development and adoption of LA tools are complex and require intense efforts in terms of time and expertise. Therefore, learning institutions often assume a “one-size-fits-all” approach, creating a single tool that is then applied across every course, discipline, and level. There has been an increase in the offering of LA tool packages that use the same off-the-shelf algorithms for all modules, in all disciplines, and at all levels [44]. However, “trace data reflects the instructional context that generated it and validity and reliability in one context is unlikely to generalize to other contexts” [34, p.22]. Gašević *et al.* [45] study demonstrates that LA predictive models must “account for instructional conditions,” as generalized models are far less powerful than course-specific models to guide practice and research. Joksimović *et al.* [46] literature review on LA approaches in massive open online courses highlights the lack of generalizability of these studies, as they adopt a widely different range of metrics to model learning. They suggest that a shared conceptualization of engagement by finding generalizable predictors could make results from future research more comparable across different contexts, and they invite a shift from observation to experimental approaches.

2.7. Insufficient Evidence of Effectiveness

Reviews of the literature highlight the lack of rigorous evaluations of the effects of LA tools [47]. Bodily and Verbert [47] literature review on student-facing LADs shows that more research is needed to understand the impact of LADs on student behavior, achievement, and skills, as the studies conducted are few and yielded mixed results. They encourage the adoption of more robust research methodologies, such as quasi-experimental studies and propensity score matching, and the investigation of underdeveloped topics, such as how students engage and interact with LADs and the evaluation of their effectiveness. Quantitative findings supporting the positive effects of LADs on learning outcomes are starting to emerge [15]; however, most of the literature consists of studies that tend to consider few outcome measures and to evaluate usability aspects, using small samples and adopting mainly qualitative strategies of inquiry [15,47]. Jivet *et al.* [9] advise that, in evaluating LA solutions, usability studies should investigate the tool’s perceived ease of use and utility and how users interpret and understand the outputs they receive. However, these aims must remain secondary to the assessment of whether the intended outcomes were achieved by LA and to the evaluation of their affective and motivational effects. The authors suggest strengthening the evaluation of LA through the triangulation of data from validated self-reported measures, assessments, and tracked data. When assessing the effectiveness of LA systems, it is essential to consider not only the outcomes but also the learning process itself. As will be explored below, researchers may use diversified data types collected by the LA system to offer valuable insights into learner activities, such as video logs, fine-grained click streams, eye-tracking data, and log files. These types of data allow for the extraction of meaningful patterns and features that can help understand learners’ intermediate states of learning and how they are related to the learning outcomes [48].

2.8. Insufficient Teacher Involvement

Teachers are one of the most important stakeholders for the integration of LA systems in schools, thus the effectiveness of LA systems is very much dependent on the acceptance and involvement of teachers [49]. As mentioned above, although teachers usually hold positive attitudes towards LA [24], they are also identified as a potential source of resistance in the adoption of these new systems [50,51]. Surveys reveal that in 2016, both in Australia and in the UK, LA initiatives were driven mostly by

IT experts and a few dedicated faculty members, but for the most part, teachers were left “out of the loop” of these novel initiatives [44]. However, teachers may develop a negative attitude towards LA systems and be reluctant to utilize them if they perceive them as lacking usefulness or ease of use [52]. Therefore, it is vital to understand and address the needs and level of tolerance that teachers have for complex systems. Moreover, teachers represent not only the end users of LA systems but also content experts in their subjects and classrooms. As educators orchestrate the teaching and learning process, they should be called to take part in the design of the learning tools they will be expected to adopt. Involving teachers as designers in the development of LA systems would help create a bridge between data and theory through the integration of the teachers’ learning design, and aid design choices that support the usability and readability of dashboards.

3. Moving Forward in LA

The previous sections highlighted the most important gaps in present LA research. Although LA offers great potential to the educational field, clear guidelines for LAD development and strong evaluation procedures are still lacking. Scarce grounding in learning theories, lack of generalizability, and subsequent scalability challenges have generated a rather large body of literature from which it is hard to draw interpretations and conclusions on the effectiveness of LA tools. Moreover, an excessive focus on data and insufficient involvement of teachers and students in the design of these systems created dashboards that are too disconnected from the instructors’ learning designs, users’ needs, and data literacy abilities, leading to usability and interpretability challenges. In this section, we are going to present some approaches that could offer useful guidelines for the future developments of LA and enhance their implementation.

3.1. Involving Teachers as Co-Designers in LA

Human-Centered Learning Analytics (HCLA) [10] proposes to overcome some limitations of LA through the participatory design of LA tools. Engaging stakeholders as co-creators holds the potential to develop more effective tools by transforming LA from something done to learners, into something done with learners. This shift could lower ethical concerns and lead to the development of tools that better fit the needs of their users. For example, this perspective is switching the focus from relying entirely on users in data interpretation to giving them the answers they are interested in.

Dimitriadis *et al.* [53] identify three fundamental principles of HCLA. First, theoretical grounding for the design and implementation of LA. Second, intensive inter-stakeholder communication in the design process, which allows to understand the needs and values of teachers, students, and other educational decision-makers. Finding a way to hear all these voices can be challenging; to facilitate the orchestration Prestigiacomo *et al.* [54] introduce OrLA, which provides a roadmap to guide communication. Lastly, integrating LA into every phase of the learning design cycle to “support Teacher Inquiry into Student Learning and evidence-based decision-making”. More specifically, during LA design, the target of the LA tools should be derived from the learning design. Then, the implementation of the LA tools can provide useful insights to inform the orchestration of learning and the evaluation of the learning design itself. Through the participatory design and the active involvement of teachers not only as end users but as designers and content experts, HCLA could favor a scalable implementation of LA and lead to the development of instruments that fit teachers’ data abilities and needs [55].

Similar principles remain valid when broadening the discussion from LA to educational AI in general. Cardona *et al.* [56] identify three instructional loops in which cooperation between AI and teachers should always center educators: the act of teaching, the planning and evaluation of teaching, and the design and evaluation of tools for teaching and learning.

3.2. Using Natural Language to Increase Interpretability

To reduce reliance on users' data literacy for LADs interpretation and support the inference process, Alhadad [6] suggests integrating textual elements into visualizations, for example, through narrative and storytelling aspects. The incorporation of storytelling in LA visualization was introduced by Echeverria *et al.* [57]. The authors advocate for the explanatory instead of the exploratory purpose of LA: dashboards should not invite the exploration of data, but rather explain insights. They propose a Learning Design Driven Data Storytelling approach, which builds on principles from Information Visualization and Data Storytelling and, in accordance with HCLA, connects them to teachers' intentions (i.e. learning design). Contrary to traditional "one-size-fits-all" data-driven visual analytics approaches, the new method derives rules from the learning design and uses them in the construction of storytelling visual analytics. Data storytelling principles determine which visual elements should be emphasized, while the learning design determines which events should be the focus of communication.

Fernandez Nieto *et al.* [58] explored the effectiveness of three visual-narrative interfaces built on three different communication methods: visual data slices, tabular visualizations, and written reports. From interviews with educators, it emerged that different methods are more helpful for different purposes. For example, written reports were perceived as useful for teachers' reflection, but not as much to be used in students' debriefings, for which tabular visualizations were thought more appropriate. Therefore, defining the purpose of the LAD and involving stakeholders in this process seems to be fundamental for the development of effective dashboards.

In an attempt to incorporate textual elements into LADs, Ramos-Soto *et al.* [59] developed a service that uses natural language templates and data sourced from the DLE to automatically generate written reports about students' activity. According to the evaluation of an expert teacher, the system was able to generate useful and overall truthful insights, albeit with small divergences and not as complete as those that would have been derived from the data by human experts.

3.3. Using Process Data to Increase Interpretability

In recent years, with the popularity of learning systems, researchers are interested in the process data, that is, the data generated while students interact with the learning systems. In a learning system, students' interactions with the user interface, including their duration on each screen and actions such as clicking, are often logged and commonly referred to as process data [60]. However, process data encompasses more than just logged data; it broadly includes empirical data that indicates the process of working on a test item based on cognitive and non-cognitive constructs [61]. This encompasses a wide variety of data types, such as action sequences, frequencies of actions, conversations or interactions within the learning system, and even eye-tracking movements and think-aloud data. In recent years, process data has received extensive research attention within the context of educational data mining, learning analytics, and artificial intelligence. Process data serves as a valuable source of detailed information regarding students' learning process within a learning system, enabling interpretation of both cognitive and behavioral aspects of learning.

As an important aspect of process data, response time has been extensively studied and is commonly regarded as an indicator of students' behaviors and cognitive processes. For example, response time has been used to identify students presenting abnormal behaviors in assessments. Wise and Ma [62] proposed a normative threshold method that compares an examinee's response time with that of their peers to determine rapid guessers or disengaged test-takers. In addition, Rios and Guo [63] developed a mixture log-normal approach which assumes that, in the presence of low effort, a bimodal response time distribution should be observed, with the lower mode representing non-effortful responding and the upper mode indicating effortful responding. This approach employs an empirical response time distribution, fits a mixed log-normal distribution, and identifies the lowest point between the two modes as the threshold. A simpler yet effective method is to visually inspect bimodal response time distributions for a distinctive gap, which can differentiate rapid guessers from

other test-takers [64]. These methods can also be extended beyond assessment environments to infer students' motivation, engagement, and learning experiences by analyzing their time spent navigating learning systems.

In addition to response time, clickstream data recorded during test-taking experiences can provide valuable insights into behavior patterns. For example, Su and Chen [65] utilized clustering techniques to group students' clickstream data with similar behavior usage patterns. Ulitzsch *et al.* [66] considered both action sequences and timing, employing cluster edge deletion to identify distinct groups of action patterns that represent common response processes. Each pattern describes a typical response process observed among test-takers. Furthermore, Tang *et al.* [67] introduced the Model Agreement Index as a measure to quantify the typicality or atypicality of an examinee's clickstream behaviors in comparison to a sequence model of behavior. To achieve this, they trained a Long Short-Term Memory network to model student behaviors. This approach allows the model to incorporate various behavior patterns and acquire knowledge about normal behavior patterns across different test-taker archetypes and styles.

Biometric measures, such as analyzing eye movements, can also offer valuable insights into the learning and test-taking behaviors of students. The duration of eye fixation can reflect the level of attention a test-taker pays to specific words in test items, with more challenging items generally requiring longer fixation periods [68]. Pupil size can be used to indicate fatigue levels, interest in specific learning content, and the cognitive workload associated with a particular task [69]. Moreover, blink rates tend to decrease when there is a higher visual demand, indicating the reallocation of cognitive resources [70]. For instance, research studies have demonstrated that when individuals encounter unfamiliar, ambiguous, or complex items, they tend to increase their regression rate, which means they look back at previous parts of the text to reinstate or confirm their cognitive effort [71,72]. Furthermore, regression (i.e., looking back to previously seen contents) has been strongly linked to the level of effort and attention a reader devotes to a reading task, thus increased regression is associated with improved accuracy of processing the content information [73].

The intermediate states of students' problem-solving or writing processes within the learning system can also be subjected to analysis. For example, Adhikari [74] proposed several process visualization practices for writing and coding tasks in learning systems, such as the playback of typing and tracking changes in paragraphs, sentences, or lines over time. By employing these visualization practices, educators are able to directly see (1) the specific points in the process where students spent the majority of their time, (2) the distribution of time between creating the initial draft and revising and editing it, (3) the paragraphs that underwent editing and revision, and (4) the paragraphs that remained unedited. These visualizations allow educators to explore, review, and analyze students' learning processes and their approach to writing or programming. In addition, students themselves can leverage these visualizations for self-reflection, direction, and improvement. Furthermore, the temporal analysis of keystrokes and backspaces informs about engagement [75] and affective states of learners [76]. Allen *et al.* [77] encourage the exploration of additional aspects of the online language production process, such as pausing typing to check syntax or research the vocabulary.

Therefore, process data can increase the interpretability of students' learning process, and including this type of data in LA systems could lead to the generation of more interpretable insights.

Moreover, NLP techniques can be employed to analyze process data. For example, Guthrie and Chen [78] analyzed log data from an online learning platform and introduced a novel approach to modeling student interactions. They incorporated information about logged event duration to differentiate between abnormally brief events and normal or extra-long events. These new event records were treated as a form of "language," where each "word" represented a student's interaction with a specific learning module, and each "sentence" captured the entire sequence of interactions. The authors used second-order Markov chains to identify patterns in this new "language" of student interactions. By visualizing these Markov chains, the authors found the interaction states associated with either disengagement or high levels of engagement. However, LLMs have been rarely applied

for log analysis. To address this gap, Chhabra [79] experimented with several models from the BERT series in order to establish a system for automatically extracting information (i.e., the events occurring within a system) from log files. In contrast to traditional log parsing approaches that heavily relied on humans constructing regular expressions, rules, or grammars for information extraction, the proposed system significantly reduced both the time and human effort required for log analysis. This work demonstrates the potential of using LLMs to extract and analyze the logged events collected through LA systems, thereby improving the ease of interpreting students' learning process.

3.4. Using Language Models to Increase Personalization

A thematic analysis of learners' attitudes toward LADs reveals that students are interested in features that support learning opportunities: they express a wish for systems that provide everyone with the same opportunities and, at the same time, a desire for customization to deliver meaningful information. They demonstrate awareness of privacy concerns and express a preference for automated alerts over personalized messages from teachers. This might be due to the fact the latter elicits feelings of surveillance [80]. Automatically generated personalized feedback could provide the benefits of customized messages without making students feel monitored by their teachers.

A literature review on automatic feedback generation (AFG) in online learning environments [81] points to the usefulness of this technology, with about half of the studies indicating that AFG enhances student performance (50.79%) and reduces teacher effort (53.96%). The main techniques used in the generation of feedback were: comparison with a desired answer, dashboards, and NLP.

NLP analyzes language in its multi-dimensionality and delivers insights about both texts and learners. Descriptive features of language (e.g., number or frequency of textual elements) can inform about student engagement, be used to predict task completion, or identify comparable texts. Characteristics of the lexicon employed in a text can be used to classify genres or estimate readability. The syntactical structure of sentences informs about the readability, quality, and complexity of the utterance, and can be used to evaluate linguistic development. Semantic analyses can identify the main message of the text and its affective connotations, or detect overlap between two texts (e.g., original text and summary). NLP analyses can also estimate cohesion and coherence, which inform about how knowledge is being processed and elaborated by learners. Moreover, NLP can communicate with teachers and learners through natural language, for example, through the generation of reports or personalized feedback [77].

Cavalcanti *et al.* [81] notice how existing studies on AFG are plagued by two of the limitations that have already been highlighted in this paper: insufficient grounding in educational theories for effective feedback, and a lack of consideration for the role of teachers in the provision of feedback. Therefore, they encourage further research for the evaluation of feedback quality and the development of tools focused on instructors. Moreover, they call for studies on the generalizability of systems for AFG, identifying a possible solution in natural language generation.

LLMs are advanced NLP models that make use of deep learning techniques to learn patterns and associations between the elements of natural language and capture statistical and contextual information from the training data. The models are usually trained on vast databases that encompass a variety of textual data sources, such as books, articles, and web pages. LLMs are not only able to understand language but also to produce coherent human-like utterances in response to any user-generated prompt. LLMs can translate, summarize and paraphrase a given text, and generate new ones. With the release of ChatGPT in November 2022, LLMs gained huge traction in society and across numerous fields, from medicine to education, as scholars explore the applications of these new systems and warn about their pitfalls. In fact, even though the training corpora is massive, it is not always accurate or up to date, which means that sometimes outputs generated by ChatGPT can be inaccurate or outdated. For example, there are records of LLMs providing links to unrelated sources or citing nonexistent literature [82]. These are examples of hallucinations, which are only one

of the unresolved challenges in LLM research [83]. Moreover, LLMs are not (yet) great at solving math problems [84].

Lim *et al.* [85] invite researchers to develop LA systems that are able to make feedback more dialogue-based; personalized feedback messages should go a step further and include comments on learning strategies (i.e. metacognitive prompts) to support sense-making, as understanding feedback and interpreting it in relation to one's own learning process is necessary to plan appropriate action in response to the feedback.

Dai *et al.* [86] provided ChatGPT with a rubric and asked it to produce feedback on student assignments, to then compare it against instructor-generated feedback. The AI tool produced fluent and coherent feedback, which received a higher average readability rating than the ones written by the teacher. Agreement between the instructor and ChatGPT was high on the evaluation of the topic of the assignment; however, precision was not as satisfactory on the evaluation of other aspects of the rubric (goal and benefit). ChatGPT generated task-focused feedback for all the students and was able to provide process-focused feedback for just over half of the assignments. On the other hand, the AI never gave feedback on self-regulation and self, while the instructor provided similar feedback in 11% and 24% of cases, respectively.

Yildirim-Erbasli and Bulut [87] discussed the potential of conversational agents in improving students' learning and assessment experiences through continuous and interactive conversations. The authors argue that conversational agents can create an interactive and dynamic learning and assessment system by administering tasks or items and offering feedback to students. The use of NLP enables conversational agents to provide real-time feedback that adapts to students' responses and their needs, fostering a more effective and engaging learning environment. Consequently, students' motivation and engagement levels in learning and test-taking can be continuously boosted through personalized conversations and directed feedback.

Another aspect of the personalization potential of LLMs is that they can be utilized to generate learning tasks or assessment items that are optimally tailored to individual student abilities. For example, LLMs have been employed to automatically generate a variety of learning and assessment materials, including reading passages [88,89], programming exercises [90], question stems [91,92], and distractors [93,94]. These examples demonstrate the potential of LLMs to create large item banks. The automatically generated assessment items then can be integrated into the existing framework of computerized adaptive tests, a testing methodology that adapts the selection of the next item based on the student's ability level inferred from their previous responses [95]. As a result, students can engage in a personalized and adaptive learning experience, thereby enhancing their engagement and improving learning outcomes [96].

With a large item bank or a bank of learning tasks created, LLMs can be further used to build recommender systems. Recommender systems in education aim to offer personalized items that match individual student preferences, needs, or ability levels, helping them navigate through educational materials and optimize their learning outcomes. Typically, there are two popular approaches for building recommender systems: collaborative filtering and content-based filtering. The underlying idea of collaborative filtering is to analyze students' past behavior and preferences to generate recommendations, identifying patterns and similarities between users or items. It assumes that students who have exhibited similar interests in the past will continue to do so in the future. On the other hand, content-based filtering examines the content of the items and compares them to students' profiles or past interactions. By identifying similarities between the content of items and students' preferences, needs, or ability levels, the system can generate recommendations that match students. In the era of LLMs, language model recommender systems have been proposed to increase transparency and control for students by enabling them to interact with the learning system using natural language [97]. LLMs are able to interpret natural language user profiles and use them to modulate learning materials for each session [98]. For example, Zhang *et al.* [99] proposed a language model recommender system leveraging several language models including GPT2 and BERT. They

converted the user-system interaction logs (items watched: 1193, 661, 914) to text inquiry ("the user watched <item name> of 1193, <item name> of 661, and <item name> of 914") and then used language models to fill in the masks for recommendation ("now the next item the user wants to watch is ____").

Therefore, LLMs can support educators in the provision of timely and personalized feedback, and give them the opportunity to devote more time to other aspects of teaching. However, it's always important to keep teachers in the loop in the process of the generation and provision of automatic feedback, as these systems are not (yet) able to touch upon all the dimensions of learning and might not integrate the learning design or take the student history into account. Some existing LA systems offer an instructor-mediated approach to personalized feedback, which offers teachers greater control over the metrics and messages returned to students, for example, by allowing them to set up "if-then" rules for message delivery based on their specific learning design. A focus group exploring students' perception of a similar system reveals that, even if they knew that the messages were to some extent automated, pupils perceived that their instructor cared about their learning. The authors argue that the perception of interpersonal communication favored proactive recipients of feedback and increased motivation for learning [100]. Cardona *et al.* [56] support the use of AI for AFG, but recommend always keeping educators at the center of the feedback loops, and invite researchers to create feedback that is not solely deficit-focused, but also asset-oriented, able to help students recognize their strengths and build onto them.

3.5. Using Language Models to Support Teachers

Feedback generation is only one of the many possible applications of LLMs to support educational practices. Allen *et al.* [77] suggest that when applying NLP to LA, we should consider both the multi-dimensional nature of language and the multiple ways in which language is part of the learning process. In fact, language permeates every aspect of learning: it is through processing natural language that learners are asked to understand course materials and tasks (input), explain their reasoning (process), and formulate their responses (output). NLP can be leveraged to analyze the learning process in all its different phases. At the input level, NLP can inform teachers about how their communications and the materials they select impact students, and also identify the most appropriate materials for each student based on their reading abilities and vocabulary skills. To understand cognitive processes underlying learning, NLP techniques can be used to automate the analysis of think-aloud protocols and open-ended questions in which respondents describe their reasoning. Lastly, NLP can analyze textual outputs produced by students with different objectives, such as automated essay scoring (AES), assessing students' abilities (e.g., vocabulary skills) and understanding of the course content, and providing highly personalized feedback.

Bonner *et al.* [101] provide examples of practical uses of LLMs to alleviate teachers' workload and free up time to focus on learners while creating engaging lessons and personalized materials. LLMs such as ChatGPT can correct grammar and evaluate cohesion in student-generated texts, summarize texts, generate presentation notes from a script, offer ideas for lessons and classroom activities, create prompts for writing exercises, generate test items, write or modify existing texts into suitable assessment materials based on skill level, and guide teachers in the development of teaching objectives and rubrics. By crafting well-thought-out and specific inputs, teachers can receive outputs that best fit their intent and meet their needs. For example, teachers can specify how many distractors to be included in the multiple choice questions generated by the AI, what writing style should be used, or how difficult the text should be. When asked to provide ideas for classroom activities to introduce a topic, ChatGPT proposed tasks that space across the taxonomy of learning, from analyzing to applying, depending on the skill level of the students that the activity was thought for. Through the use of LLMs, teachers can create personalized materials for each student in a fraction of the time it would take them to do so themselves.

AI technologies could enhance the practices of formative assessment by capturing complex competencies, such as teamwork and self-regulation, by promoting accessibility for neurodivergent

learners, or by offering students constant support whenever needed, even outside of class times [56]. For example, LLMs can be used to build virtual tutors that can help learners understand concepts, test their knowledge, improve their writing, or solve assignments. Khanmigo is a virtual tutor developed by Khan Academy that uses GPT-4 to support both students and teachers in many of the ways presented above. The system was instructed to tutor students based on the best practices identified by the literature, which means it supports and guides student reasoning processes without doing the assignment for them, even when asked to do so. Chat logs are made available for teachers to access, and inappropriate requests (e.g., cheating) are automatically flagged by the system and brought to the educator's attention [102].

All these applications are anticipated to reduce teachers' workload, either by taking it on themselves (e.g., modifying a text so that it meets the appropriate difficulty level for learners) or by offering educators guidance and ideas (e.g., planning classroom activities). Users are encouraged to be specific when providing prompts and to keep interacting with the LLMs giving them further instructions if they are not satisfied with the answer they received, as these systems retain a more or less extensive memory of the conversation (context window). Increasing efforts have been recently focused on enlarging the mnemonic capabilities of LLMs, which would be really useful to approach complex tasks, such as summarizing entire books, or keeping the memory of each student's background and interests.

4. Discussion

Although the literature has received numerous contributions over the last few years, there are still limitations in the development and design of LA tools and challenges in their implementation. Existing LA applications still suffer from an insufficient grounding in pedagogical theories, leading to difficulties in the valid interpretation and use of the learner data. Moreover, generalizability is poor and substantial evidence of LA effectiveness is lacking, which is due both to mixed results and to the paucity of evaluation studies making use of strong research methodologies. Educators' overall attitude towards LA tends to be positive, but they still face challenges in the adoption of LA tools. An excessive focus on data, detached from learning theories and from the teacher's learning design, together with poor design choices, can create tools that don't meet the needs and data literacy abilities of their end users.

Making teachers co-designers in the development of LA seems to be a promising route to integrate pedagogical theories and the teachers' own learning design with the behavioral data collected in the DLE. The collaborative design process proposed by HCLA should yield tools that better meet the context need, better enable teachers to interpret insights, and better meet their data literacy skills.

Another promising way to aid users in the interpretation of LA data is to integrate visual information with written text. Natural language is central to communication, and it permeates every aspect of teaching and learning. With the recent and fast evolution of language models, a plethora of new opportunities are opening up in the educational field. LLMs can be used to evaluate assignments, provide personalized feedback on students' essays and on their progress, or offer support as an ever-accessible tutor. Furthermore, LLMs can support teachers in a variety of other tasks, from adapting learning materials to their students' language proficiency levels to developing creative activities, learning plans, essay prompts, or questions for testing. LLMs should not be embraced as the solution to all problems in education and LA; they could prove powerful in increasing interpretability and personalization of LA insights, but cannot address the foundational issues related to the development of LA systems and to the investigation of their effects.

Integrating LLMs into LA could make insights more interpretable for users, and integrating LA into LLMs could give the language model the context necessary to offer each student highly personalized and better-rounded feedback, that takes their history, progress, and interests into account when providing reports and recommendations. Moreover, LLMs can serve educators as support tools to approach complex and time-consuming teaching tasks. The intent is not to use AI to replace teachers,

but to put technology at the service of teachers. Educators should take advantage of LLMs as a resource to reduce workload, stimulate creativity and offer students tailored materials and more feedback while retaining their role of reference figures and decision-makers in the planning and evaluation of learning. AI is not supposed to strip teachers of the value of their expertise, but rather to support it and allow them to focus on tasks in which the human factor cannot be replaced.

Contrary to the interpretation of LA data, LLMs outputs are as straightforward as it gets, since the systems communicate directly through natural language. In this regard, one of the barriers to acceptance and usability is removed. However, integrating LLMs systems into teaching practices still requires trust in the technology and an adjustment in the ways teachers have been operating until now. As applications of LLMs increasingly take hold in the educational world, we should provide educators with guidelines on how to interact with these systems, including how to phrase their prompts to get the answer that best fits their needs, what are the limitations of these tools and what are the risks to be mindful about. When used responsibly, LLMs such as ChatGPT present opportunities to enhance students' learning experience and mitigate a considerable amount of workload for teachers, for example, through assistance in the formulation of test item writing [103]. However, educators should be aware of potential issues that LLMs entail, such as over-reliance on the LLM, copyright, and cheating [103,104]. In this era of rapid technological development, a new approach to teaching practices may be necessary to revolutionize modern education and reconcile the tension between human teachers and artificial intelligence [104]. In particular, to successfully establish a safe and prolific cooperation with AI in education, we need to find a balance between the contrasting forces of human control and delegation to technology, between collecting more data to better represent students and respecting their privacy, and strive for personalization that does not cross over the line of teacher surveillance [56].

4.1. Limitations and Directions for Future Research

This study has several limitations worth noting. First, we recognize that this paper does not offer any AI- or LA-based solutions to overcome the limitations in the evaluation stage of the LA life cycle. Specifically, the issue of generalizability is an ongoing challenge for LA researchers. Inadequate feature representation, inadequate sample size, and imbalanced class are primary causes that hinder the generalizability of LA models [105]. However, such problems are commonly encountered in real-world datasets. In fact, the achievement of a shared conceptualization is hampered by patterns in the population, as both individual factors (e.g., the shift in interest) and societal factors (e.g., trends in education) could change at the sub-group level and therefore hinder a common feature representation. The mentioned sample size and imbalanced class issues are also hardly avoidable, as in predictive tasks that target low-occurrence but high-impact situations, such as school dropout, the discrepancy between the minority and the majority class is usually high [106], and thus these limitations can usually be addressed only after the fact.

Furthermore, the challenge of insufficient evidence of effectiveness too cannot be addressed solely by using AI- or LA-based solutions, but it calls for purposeful choices in the development of LA tools and evaluation studies. To guide the planning of LA evaluations, we encourage future research to follow Jivet *et al.* [9]'s recommendations outlined above. However, future evaluation studies might employ NLP techniques and LLMs to support the qualitative analysis of teachers' and students' responses to open-ended questions about LA usability and perceived utility.

Lastly, the LLM-based educational tools discussed above are still under development, and technological readiness represents a challenge to their promised benefits. Yan *et al.* [107] conducted a scoping review on the applications of LLMs in educational tasks, focusing on the practical and ethical limitations of LLM applications. The authors asserted that there was little evidence for the successful implementation of LLM-based innovations in real educational practices. In addition, they noted that existing LLMs applications are still in the early stages of technology readiness and struggle to handle complex educational tasks effectively, despite showing high performance in simple tasks

like sentiment analysis of student feedback [108]. Furthermore, the authors pointed out that many reviewed studies lacked sufficient details about their methodologies (e.g., not open-sourcing the data and/or codes used for analysis), making it challenging for other researchers and practitioners to replicate their proposed LLMs-based innovations. Based on the results, Yan *et al.* [107] suggest future studies to validate LLM-based education technologies through their deployment and integration in real classrooms and educational settings. Real-world studies would allow researchers to test the models' performance in authentic scenarios, particularly for tasks of prediction and generation, and to evaluate their generalizability. The authors warn researchers that studies in educational technology tend to suffer from limited replicability, therefore they encourage them to open-source their models and share enough details about their datasets.

From an ethical perspective, the adoption of LLMs and AI-powered learning technologies in education should carefully consider their accountability, explainability, fairness, interpretability, and safety [109]. The majority of existing LLMs-based innovations are considered transparent and understandable only by AI researchers and practitioners, while none are perceived as sufficiently transparent by educational stakeholders, such as teachers and students [107]. In order to address this issue, future research should incorporate a human-in-the-loop component, actively involving educational stakeholders in the development and evaluation process. This also ensures that the educational stakeholders gain insights into how LLMs and AI-powered learning technologies function and how they can be harnessed effectively for improved learning outcomes.

Future studies could further explore the application of NLP techniques to analyze process data and generate written reports from students' data. Although LLMs have only recently been developed and numerous challenges remain to be solved, researchers both inside and outside of academia are hastily at work to address them and improve these models; and as the capabilities of LLMs expand, so will their applications [83]. For example, expanding LLMs context windows would support the provision of feedback that takes into account students' background information, such as individual interests and level of language proficiency [110]. Further, LLM could also be used with an intelligent tutoring system to enhance the quality of feedback provided to students [111]. Moreover, as LLMs will find their way into teaching and learning practices, further consideration should be given to the ethical implications of AI in education. Data privacy and transparency concerns call for higher model explainability and greater involvement of stakeholders in the development and evaluation of educational technologies. Moreover, while the high level of personalization that LLMs could offer students might increase equity, the costs currently associated with developing and adopting these technologies raise issues about equality. Additional concerns involve model accuracy, discrimination, and bias [107]. Researchers, policymakers, and other educational stakeholders should all think about what they can do to mitigate these threats to fairness and ensure that educational AI will not broaden inequalities instead of reducing them.

5. Conclusions

While LA holds many promises to enhance teaching and learning, there is still work to be done to bring them to full fruition. The present paper highlighted the main limitations in the development, implementation, and evaluation of LA, and offered guidelines and ideas to overcome some of these challenges. In particular, there is a need for incorporating data and learning theories, as these would provide a lens to make sense of LA insights. HCLA offers principles to reach this integration, through intensive cooperation with educators as co-designers of LA solutions. In addition, using process data in LA systems can enhance our understanding of students' learning processes and increase the interpretability of insights. Furthermore, we explored numerous ways in which LLMs can be deployed to make LA insights more interpretable and customizable, to increase personalization through feedback generation and content recommendation, and to support teachers' tasks more broadly while always maintaining a human-centered approach.

References

1. Society for Learning Analytics Research [SoLAR]. What is Learning Analytics?, 2021.
2. Siemens, G. Learning analytics: envisioning a research discipline and a domain of practice. In Proceedings of the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge; Association for Computing Machinery: New York, NY, USA, 2012; LAK '12, pp. 4–8. <https://doi.org/10.1145/2330601.2330605>.
3. Lee, L.K.; Cheung, S.K.S.; Kwok, L.F. Learning analytics: current trends and innovative practices. *Journal of Computers in Education* **2020**, *7*, 1–6. <https://doi.org/10.1007/s40692-020-00155-8>.
4. Matcha, W.; Uzir, N.A.; Gašević, D.; Pardo, A. A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective. *IEEE Transactions on Learning Technologies* **2020**, *13*, 226–245. Conference Name: IEEE Transactions on Learning Technologies, <https://doi.org/10.1109/TLT.2019.2916802>.
5. Wang, Q.; Mousavi, A.; Lu, C. A scoping review of empirical studies on theory-driven learning analytics. *Distance Education* **2022**, *43*, 6–29. Publisher: Routledge _eprint: <https://doi.org/10.1080/01587919.2021.2020621>, <https://doi.org/10.1080/01587919.2021.2020621>.
6. Alhadad, S.S.J. Visualizing Data to Support Judgement, Inference, and Decision Making in Learning Analytics: Insights from Cognitive Psychology and Visualization Science. *Journal of Learning Analytics* **2018**, *5*, 60–85. Number: 2, <https://doi.org/10.18608/jla.2018.52.5>.
7. Wong, B.T.m.; Li, K.C. A review of learning analytics intervention in higher education (2011–2018). *Journal of Computers in Education* **2020**, *7*, 7–28. <https://doi.org/10.1007/s40692-019-00143-7>.
8. Bodily, R.; Ikahihifo, T.K.; Mackley, B.; Graham, C.R. The design, development, and implementation of student-facing learning analytics dashboards. *Journal of Computing in Higher Education* **2018**, *30*, 572–598. <https://doi.org/10.1007/s12528-018-9186-0>.
9. Jivet, I.; Scheffel, M.; Specht, M.; Drachler, H. License to evaluate: preparing learning analytics dashboards for educational practice. In Proceedings of the Proceedings of the 8th International Conference on Learning Analytics and Knowledge; ACM: Sydney New South Wales Australia, 2018; pp. 31–40. <https://doi.org/10.1145/3170358.3170421>.
10. Buckingham Shum, S.; Ferguson, R.; Martinez-Maldonado, R. Human-Centred Learning Analytics. *Journal of Learning Analytics* **2019**, *6*, 1–9. Number: 2.
11. UNESCO. Section 2: preparing learners to thrive in the future with AI. In *Artificial intelligence in education: Challenges and opportunities for sustainable development*; Working Papers on Education Policy, the United Nations Educational, Scientific and Cultural Organization.: France, 2019; pp. 17–24.
12. Schwendimann, B.A.; Rodríguez-Triana, M.J.; Vozniuk, A.; Prieto, L.P.; Boroujeni, M.S.; Holzer, A.; Gillet, D.; Dillenbourg, P. Perceiving Learning at a Glance: A Systematic Literature Review of Learning Dashboard Research. *IEEE Transactions on Learning Technologies* **2017**, *10*, 30–41. Conference Name: IEEE Transactions on Learning Technologies, <https://doi.org/10.1109/TLT.2016.2599522>.
13. Sedrakyan, G.; Malmberg, J.; Verbert, K.; Järvelä, S.; Kirschner, P.A. Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. *Computers in Human Behavior* **2020**, *107*, 105512. <https://doi.org/10.1016/j.chb.2018.05.004>.
14. Rets, I.; Herodotou, C.; Bayer, V.; Hlosta, M.; Rienties, B. Exploring critical factors of the perceived usefulness of a learning analytics dashboard for distance university students. *International Journal of Educational Technology in Higher Education* **2021**, *18*, 46. <https://doi.org/10.1186/s41239-021-00284-9>.
15. Susnjak, T.; Ramaswami, G.S.; Mathrani, A. Learning analytics dashboard: a tool for providing actionable insights to learners. *International Journal of Educational Technology in Higher Education* **2022**, *19*, 12. <https://doi.org/10.1186/s41239-021-00313-7>.
16. Valle, N.; Antonenko, P.; Valle, D.; Sommer, M.; Huggins-Manley, A.C.; Dawson, K.; Kim, D.; Baiser, B. Predict or describe? How learning analytics dashboard design influences motivation and statistics anxiety in an online statistics course. *Educational Technology Research and Development* **2021**, *69*, 1405–1431. <https://doi.org/10.1007/s11423-021-09998-z>.
17. Iraj, H.; Fudge, A.; Khan, H.; Faulkner, M.; Pardo, A.; Kovanović, V. Narrowing the Feedback Gap: Examining Student Engagement with Personalized and Actionable Feedback Messages. *Journal of Learning Analytics* **2021**, *8*, 101–116. Number: 3, <https://doi.org/10.18608/jla.2021.7184>.

18. Wagner, E.; Ice, P. Data changes everything: Delivering on the promise of learning analytics in higher education. *Educause Review* **2012**, *47*, 32.
19. Algayres, M.G.; Triantafyllou, E. Learning Analytics in Flipped Classrooms: a Scoping Review. *Electronic Journal of E-Learning* **2020**, *18*, 397–409. <https://doi.org/10.34190/JEL.18.5.003>.
20. Wise, A.F.; Shaffer, D.W. Why Theory Matters More than Ever in the Age of Big Data. *Journal of Learning Analytics* **2015**, *2*, 5–13. Number: 2, <https://doi.org/10.18608/jla.2015.22.2>.
21. You, J.W. Identifying significant indicators using LMS data to predict course achievement in online learning. *The Internet and Higher Education* **2016**, *29*, 23–30. <https://doi.org/10.1016/j.iheduc.2015.11.003>.
22. Caspari-Sadeghi, S. Applying Learning Analytics in Online Environments: Measuring Learners' Engagement Unobtrusively. *Frontiers in Education* **2022**, *7*.
23. Few, S. Dashboard Design: Taking a Metaphor Too Far. *DM Review* **2005**, *15*, 18. Num Pages: 0 Place: New York, United States Publisher: SourceMedia Section: Data Visualization.
24. McKenney, S.; Mor, Y. Supporting teachers in data-informed educational design. *British Journal of Educational Technology* **2015**, *46*, 265–279. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/bjet.12262>, <https://doi.org/10.1111/bjet.12262>.
25. van Leeuwen, A. Teachers' perceptions of the usability of learning analytics reports in a flipped university course: when and how does information become actionable knowledge? *Educational Technology Research and Development* **2019**, *67*, 1043–1064. <https://doi.org/10.1007/s11423-018-09639-y>.
26. Davis, F.D.; Bagozzi, R.P.; Warshaw, P.R. User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science* **1989**, *35*, 982–1003. Publisher: INFORMS, <https://doi.org/10.1287/mnsc.35.8.982>.
27. van Leeuwen, A.; Janssen, J.; Erkens, G.; Brekelmans, M. Supporting teachers in guiding collaborating students: Effects of learning analytics in CSCL. *Computers & Education* **2014**, *79*, 28–39. <https://doi.org/10.1016/j.compedu.2014.07.007>.
28. van Leeuwen, A.; Janssen, J.; Erkens, G.; Brekelmans, M. Teacher regulation of cognitive activities during student collaboration: Effects of learning analytics. *Computers & Education* **2015**, *90*, 80–94. <https://doi.org/10.1016/j.compedu.2015.09.006>.
29. Ramaswami, G.; Susnjak, T.; Mathrani, A.; Umer, R. Use of Predictive Analytics within Learning Analytics Dashboards: A Review of Case Studies. *Technology, Knowledge and Learning* **2022**. <https://doi.org/10.1007/s10758-022-09613-x>.
30. Liu, R.; Koedinger, K.R. Closing the Loop: Automated Data-Driven Cognitive Model Discoveries Lead to Improved Instruction and Learning Gains. *Journal of Educational Data Mining* **2017**, *9*, 25–41. <https://doi.org/10.5281/zenodo.3554626>.
31. Bañeres, D.; Rodríguez, M.E.; Guerrero-Roldán, A.E.; Karadeniz, A. An Early Warning System to Detect At-Risk Students in Online Higher Education. *Applied Sciences* **2020**, *10*, 4427. Number: 13 Publisher: Multidisciplinary Digital Publishing Institute, <https://doi.org/10.3390/app10134427>.
32. Namoun, A.; Alshantiti, A. Predicting Student Performance Using Data Mining and Learning Analytics Techniques: A Systematic Literature Review. *Applied Sciences* **2021**, *11*, 237. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute, <https://doi.org/10.3390/app11010237>.
33. Jayaprakash, S.M.; Moody, E.W.; Lauría, E.J.M.; Regan, J.R.; Baron, J.D. Early Alert of Academically At-Risk Students: An Open Source Analytics Initiative. *Journal of Learning Analytics* **2014**, *1*, 6–47. Number: 1, <https://doi.org/10.18608/jla.2014.11.3>.
34. Gray, G.; Bergner, Y. A Practitioner's Guide to Measurement in Learning Analytics: Decisions, Opportunities, and Challenges. In *The Handbook of Learning Analytics*, 2 ed.; Lang, C.; Siemens, G.; Wise, A.F.; GaÅjeviÄŹ, D.; Mercerón, A., Eds.; SoLAR, 2022; pp. 20–28. Section: 2.
35. Hattie, J.; Timperley, H. The power of feedback. *Review of educational research* **2007**, *77*, 81–112.
36. Carless, D.; Boud, D. The development of student feedback literacy: enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, *43*, 1315–1325. Publisher: Routledge _eprint: <https://doi.org/10.1080/02602938.2018.1463354>, <https://doi.org/10.1080/02602938.2018.1463354>.
37. Sutton, P. Conceptualizing feedback literacy: knowing, being, and acting. *Innovations in Education and Teaching International* **2012**, *49*, 31–40. Publisher: Routledge _eprint: <https://doi.org/10.1080/14703297.2012.647781>, <https://doi.org/10.1080/14703297.2012.647781>.

38. Irons, A. *Enhancing Learning through Formative Assessment and Feedback*; Routledge: London, 2007. <https://doi.org/10.4324/9780203934333>.
39. Karaoglan Yilmaz, F.G.; Yilmaz, R. Learning analytics as a metacognitive tool to influence learner transactional distance and motivation in online learning environments. *Innovations in Education and Teaching International* **2021**, *58*, 575–585. Publisher: Routledge _eprint: <https://doi.org/10.1080/14703297.2020.1794928>, <https://doi.org/10.1080/14703297.2020.1794928>.
40. Butler, D.L.; Winne, P.H. Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research* **1995**, *65*, 245–281. Publisher: American Educational Research Association, <https://doi.org/10.3102/00346543065003245>.
41. Dawson, P.; Henderson, M.; Ryan, T.; Mahoney, P.; Boud, D.; Phillips, M.; Molloy, E. Technology and Feedback Design. In *Learning, Design, and Technology: An International Compendium of Theory, Research, Practice, and Policy*; Spector, M.J.; Lockee, B.B.; Childress, M.D., Eds.; Springer International Publishing: Cham, 2018; pp. 1–45. https://doi.org/10.1007/978-3-319-17727-4_124-1.
42. Pardo, A.; Jovanovic, J.; Dawson, S.; Gašević, D.; Mirriahi, N. Using learning analytics to scale the provision of personalised feedback. *British Journal of Educational Technology* **2019**, *50*, 128–138. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/bjet.12592>, <https://doi.org/10.1111/bjet.12592>.
43. Evans, C. Making Sense of Assessment Feedback in Higher Education. *Review of Educational Research* **2013**, *83*, 70–120. Publisher: American Educational Research Association, <https://doi.org/10.3102/0034654312474350>.
44. Wilson, A.; Watson, C.; Thompson, T.L.; Drew, V.; Doyle, S. Learning analytics: challenges and limitations. *Teaching in Higher Education* **2017**, *22*, 991–1007. Publisher: Routledge _eprint: <https://doi.org/10.1080/13562517.2017.1332026>, <https://doi.org/10.1080/13562517.2017.1332026>.
45. Gašević, D.; Dawson, S.; Rogers, T.; Gasevic, D. Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education* **2016**, *28*, 68–84. <https://doi.org/10.1016/j.iheduc.2015.10.002>.
46. Joksimović, S.; Poquet, O.; Kovanović, V.; Dowell, N.; Mills, C.; Gašević, D.; Dawson, S.; Graesser, A.C.; Brooks, C. How Do We Model Learning at Scale? A Systematic Review of Research on MOOCs. *Review of Educational Research* **2018**, *88*, 43–86. Publisher: American Educational Research Association, <https://doi.org/10.3102/0034654317740335>.
47. Bodily, R.; Verbert, K. Review of Research on Student-Facing Learning Analytics Dashboards and Educational Recommender Systems. *IEEE Transactions on Learning Technologies* **2017**, *10*, 405–418. Conference Name: IEEE Transactions on Learning Technologies, <https://doi.org/10.1109/TLT.2017.2740172>.
48. Greer, J.; Mark, M. Evaluation methods for intelligent tutoring systems revisited. *International Journal of Artificial Intelligence in Education* **2016**, *26*, 387–392.
49. Islahi, F.; Nasrin. Exploring Teacher Attitude towards Information Technology with a Gender Perspective. *Contemporary Educational Technology* **2019**, *10*, 37–54. Publisher: Bastas, <https://doi.org/10.30935/cet.512527>.
50. Herodotou, C.; Hlostá, M.; Boroowa, A.; Rienties, B.; Zdrahal, Z.; Mangafa, C. Empowering online teachers through predictive learning analytics. *British Journal of Educational Technology* **2019**, *50*, 3064–3079. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/bjet.12853>, <https://doi.org/10.1111/bjet.12853>.
51. Herodotou, C.; Rienties, B.; Boroowa, A.; Zdrahal, Z.; Hlostá, M. A large-scale implementation of predictive learning analytics in higher education: the teachers' role and perspective. *Educational Technology Research and Development* **2019**, *67*, 1273–1306. <https://doi.org/10.1007/s11423-019-09685-0>.
52. Sabraz Nawaz, S.; Thowfeek, M.H.; Rashida, M.F. School Teachers' intention to use E-Learning systems in Sri Lanka: A modified TAM approach. *Information and Knowledge Management* **2015**.
53. Dimitriadis, Y.; Martínez-Maldonado, R.; Wiley, K. Human-Centered Design Principles for Actionable Learning Analytics. In *Research on E-Learning and ICT in Education: Technological, Pedagogical and Instructional Perspectives*; Tsiatsos, T.; Demetriadis, S.; Mikropoulos, A.; Dagdilelis, V., Eds.; Springer International Publishing: Cham, 2021; pp. 277–296. https://doi.org/10.1007/978-3-030-64363-8_15.
54. Prestigiacomo, R.; Hadgraft, R.; Hunter, J.; Locker, L.; Knight, S.; Van Den Hoven, E.; Martinez-Maldonado, R. Learning-centred translucence: an approach to understand how teachers talk about classroom data. *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* **2020**, pp. 100–105. Conference Name: LAK '20: 10th International Conference on Learning Analytics and Knowledge ISBN: 9781450377126 Place: Frankfurt Germany Publisher: ACM, <https://doi.org/10.1145/3375462.3375475>.

55. Herodotou, C.; Rienties, B.; Hlosta, M.; Boroowa, A.; Mangafa, C.; Zdrahal, Z. The scalable implementation of predictive learning analytics at a distance learning university: Insights from a longitudinal case study. *The Internet and Higher Education* **2020**, *45*, 100725. <https://doi.org/10.1016/j.iheduc.2020.100725>.
56. Cardona, M.A.; Rodríguez, R.J.; Ishmael, K. Artificial Intelligence and Future of Teaching and Learning: Insights and Recommendations. Technical report, US Department of Education, Office of Educational Technology, Washington DC, 2023.
57. Echeverria, V.; Martinez-Maldonado, R.; Shum, S.B.; Chiluiza, K.; Granda, R.; Conati, C. Exploratory versus Explanatory Visual Learning Analytics: Driving Teachers' Attention through Educational Data Storytelling. *Journal of Learning Analytics* **2018**, *5*, 73–97–73–97. Number: 3, <https://doi.org/10.18608/jla.2018.53.6>.
58. Fernandez Nieto, G.M.; Kitto, K.; Buckingham Shum, S.; Martinez-Maldonado, R. Beyond the Learning Analytics Dashboard: Alternative Ways to Communicate Student Data Insights Combining Visualisation, Narrative and Storytelling. In Proceedings of the LAK22: 12th International Learning Analytics and Knowledge Conference; Association for Computing Machinery: New York, NY, USA, 2022; LAK22, pp. 219–229. <https://doi.org/10.1145/3506860.3506895>.
59. Ramos-Soto, A.; Vazquez-Barreiros, B.; Bugarín, A.; Gewerc, A.; Barro, S. Evaluation of a Data-To-Text System for Verbalizing a Learning Analytics Dashboard. *International Journal of Intelligent Systems* **2017**, *32*, 177–193. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/int.21835>, <https://doi.org/10.1002/int.21835>.
60. He, Q.; von Davier, M. Identifying Feature Sequences from Process Data in Problem-Solving Items with N-Grams. In Proceedings of the Quantitative Psychology Research; van der Ark, L.A.; Bolt, D.M.; Wang, W.C.; Douglas, J.A.; Chow, S.M., Eds.; Springer International Publishing: Cham, 2015; Springer Proceedings in Mathematics & Statistics, pp. 173–190. https://doi.org/10.1007/978-3-319-19977-1_13.
61. Reis Costa, D.; Leoncio Netto, W. Process Data Analysis in ILSAs. In *International Handbook of Comparative Large-Scale Studies in Education: Perspectives, Methods and Findings*; Nilsen, T.; Stancel-Piatak, A.; Gustafsson, J.E., Eds.; Springer International Handbooks of Education, Springer International Publishing: Cham, 2022; pp. 1–27. https://doi.org/10.1007/978-3-030-38298-8_60-1.
62. Wise, S.L.; Ma, L. Setting response time thresholds for a CAT item pool: The normative threshold method. In Proceedings of the annual meeting of the National Council on Measurement in Education, Vancouver, Canada, 2012, pp. 163–183.
63. Rios, J.A.; Guo, H. Can Culture Be a Salient Predictor of Test-Taking Engagement? An Analysis of Differential Noneffortful Responding on an International College-Level Assessment of Critical Thinking. *Applied Measurement in Education* **2020**, *33*, 263–279. <https://doi.org/10.1080/08957347.2020.1789141>.
64. Wise, S.L.; Kong, X. Response Time Effort: A New Measure of Examinee Motivation in Computer-Based Tests. *Applied Measurement in Education* **2005**, *18*, 163–183. Publisher: Routledge _eprint: https://doi.org/10.1207/s15324818ame1802_2, https://doi.org/10.1207/s15324818ame1802_2.
65. Su, Q.; Chen, L. A method for discovering clusters of e-commerce interest patterns using click-stream data. *Electronic Commerce Research and Applications* **2015**, *14*, 1–13. <https://doi.org/10.1016/j.elerap.2014.10.002>.
66. Ulitzsch, E.; He, Q.; Ulitzsch, V.; Molter, H.; Nichterlein, A.; Niedermeier, R.; Pohl, S. Combining Clickstream Analyses and Graph-Modeled Data Clustering for Identifying Common Response Processes. *Psychometrika* **2021**, *86*, 190–214. <https://doi.org/10.1007/s11336-020-09743-0>.
67. Tang, S.; Samuel, S.; Li, Z. Detecting atypical test-taking behaviors with behavior prediction using LSTM. *Psychological Test and Assessment Modeling* **2023**, *65*, 76–124.
68. Rayner, K. Eye movements in reading and information processing: 20 years of research. *Psychological bulletin* **1998**, *124*, 372.
69. Morad, Y.; Lemberg, H.; Yofe, N.; Dagan, Y. Pupillography as an objective indicator of fatigue. *Current Eye Research* **2000**, *21*, 535–542.
70. Benedetto, S.; Pedrotti, M.; Minin, L.; Baccino, T.; Re, A.; Montanari, R. Driver workload and eye blink duration. *Transportation Research Part F: Traffic Psychology and Behaviour* **2011**, *14*, 199–208. <https://doi.org/10.1016/j.trf.2010.12.001>.
71. Booth, R.W.; Weger, U.W. The function of regressions in reading: Backward eye movements allow rereading. *Memory & Cognition* **2013**, *41*, 82–97. <https://doi.org/10.3758/s13421-012-0244-y>.
72. Inhoff, A.W.; Greenberg, S.N.; Solomon, M.; Wang, C.A. Word integration and regression programming during reading: a test of the E-Z reader 10 model. *Journal of Experimental Psychology. Human Perception and Performance* **2009**, *35*, 1571–1584. <https://doi.org/10.1037/a0014250>.

73. Coëffé, C.; O'regan, J.K. Reducing the influence of non-target stimuli on saccade accuracy: Predictability and latency effects. *Vision research* **1987**, *27*, 227–240.
74. Adhikari, B. Thinking beyond chatbots' threat to education: Visualizations to elucidate the writing and coding process, 2023, [arXiv:cs.CY/2304.14342].
75. Allen, L.K.; Mills, C.; Jacovina, M.E.; Crossley, S.; D'Mello, S.; McNamara, D.S. Investigating boredom and engagement during writing using multiple sources of information: the essay, the writer, and keystrokes. In Proceedings of the Proceedings of the Sixth International Conference on Learning Analytics & Knowledge; Association for Computing Machinery: New York, NY, USA, 2016; LAK '16, pp. 114–123. <https://doi.org/10.1145/2883851.2883939>.
76. Bixler, R.; D'Mello, S. Detecting boredom and engagement during writing with keystroke analysis, task appraisals, and stable traits. In Proceedings of the Proceedings of the 2013 international conference on Intelligent user interfaces; Association for Computing Machinery: New York, NY, USA, 2013; IUI '13, pp. 225–234. <https://doi.org/10.1145/2449396.2449426>.
77. Allen, L.; Creer, S.; Oncel, P. Chapter 5: Natural Language Processing as a Tool for Learning Analytics -Towards a Multi-Dimensional View of the Learning Process. In *The Handbook of Learning Analytics*; Society of Learning Analytics Research, 2022. <https://doi.org/10.18608/hla22.005>.
78. Guthrie, M.; Chen, Z. Adding duration-based quality labels to learning events for improved description of students' online learning behavior. In Proceedings of the Proceedings of the 12th International Conference on Educational Data Mining (2019), 2019.
79. Chhabra, A. A System for Automatic Information Extraction from Log Files. PhD thesis, Université d'Ottawa/University of Ottawa, 2022.
80. Roberts, L.D.; Howell, J.A.; Seaman, K. Give Me a Customizable Dashboard: Personalized Learning Analytics Dashboards in Higher Education. *Technology, Knowledge and Learning* **2017**, *22*, 317–333. <https://doi.org/10.1007/s10758-017-9316-1>.
81. Cavalcanti, A.P.; Barbosa, A.; Carvalho, R.; Freitas, F.; Tsai, Y.S.; Gašević, D.; Mello, R.F. Automatic feedback in online learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence* **2021**, *2*, 100027. <https://doi.org/10.1016/j.caeai.2021.100027>.
82. McGee, R.W. Is ESG a Bad Idea? The Chatgpt Response, 2023. <https://doi.org/10.2139/ssrn.4413431>.
83. Kaddour, J.; Harris, J.; Mozes, M.; Bradley, H.; Raileanu, R.; McHardy, R. Challenges and Applications of Large Language Models, 2023, [arXiv:cs.CL/2307.10169].
84. Frieder, S.; Pinchetti, L.; Griffiths, R.R.; Salvatori, T.; Lukasiewicz, T.; Petersen, P.C.; Chevalier, A.; Berner, J. Mathematical Capabilities of ChatGPT, 2023. arXiv:2301.13867 [cs], <https://doi.org/10.48550/arXiv.2301.13867>.
85. Lim, L.A.; Dawson, S.; Gašević, D.; Joksimović, S.; Fudge, A.; Pardo, A.; Gentili, S. Students' sense-making of personalised feedback based on learning analytics. *Australasian Journal of Educational Technology* **2020**, *36*, 15–33. Number: 6, <https://doi.org/10.14742/ajet.6370>.
86. Dai, W.; Lin, J.; Jin, F.; Li, T.; Tsai, Y.S.; Gasevic, D.; Chen, G. Can Large Language Models Provide Feedback to Students? A Case Study on ChatGPT, 2023. <https://doi.org/10.35542/osf.io/hcgzj>.
87. Yildirim-Erbasli, S.N.; Bulut, O. Conversation-based assessment: A novel approach to boosting test-taking effort in digital formative assessment. *Computers and Education: Artificial Intelligence* **2023**, *4*, 100135. <https://doi.org/https://doi.org/10.1016/j.caeai.2023.100135>.
88. Bulut, O.; Yildirim-Erbasli, S.N. Automatic story and item generation for reading comprehension assessments with transformers. *International Journal of Assessment Tools in Education* **2022**, *9*, 72–87. <https://doi.org/10.21449/ijate.1124382>.
89. Attali, Y.; Runge, A.; LaFlair, G.T.; Yancey, K.; Goodwin, S.; Park, Y.; von Davier, A.A. The interactive reading task: Transformer-based automatic item generation. *Frontiers in Artificial Intelligence* **2022**, *5*.
90. Sarsa, S.; Denny, P.; Hellas, A.; Leinonen, J. Automatic generation of programming exercises and code explanations using large language models. In Proceedings of the Proceedings of the 2022 ACM Conference on International Computing Education Research-Volume 1, 2022, pp. 27–43.
91. Tsai, D.; Chang, W.; Yang, S. Short Answer Questions Generation by Fine-Tuning BERT and GPT-2. In Proceedings of the 29th International Conference on Computers in Education Conference, ICCE 2021 - Proceedings; Rodrigo, M.; Iyer, S.; Mitrovic, A.; Cheng, H.; Kohen-Vacs, D.; Matuk, C.; Palalas, A.; Rajenran,

- R.; Seta, K.; Wang, J., Eds. Asia-Pacific Society for Computers in Education, 2021, 29th International Conference on Computers in Education Conference, ICCE 2021 - Proceedings, pp. 509–515.
92. Shan, J.; Nishihara, Y.; Yamanishi, R.; Maeda, A. Question Generation for Reading Comprehension of Language Learning Test : -A Method using Seq2Seq Approach with Transformer Model-. In Proceedings of the 2019 International Conference on Technologies and Applications of Artificial Intelligence (TAAI), 2019, pp. 1–6. <https://doi.org/10.1109/TAAI48200.2019.8959903>.
 93. Offerijns, J.; Verberne, S.; Verhoef, T. Better Distractions: Transformer-based Distractor Generation and Multiple Choice Question Filtering, 2020. arXiv:2010.09598 [cs], <https://doi.org/10.48550/arXiv.2010.09598>.
 94. Zu, J.; Choi, I.; Hao, J. Automated distractor generation for fill-in-the-blank items using a prompt-based learning approach. *Psychological Testing and Assessment Modeling* **2023**, 65, 55–75.
 95. Wainer, H.; Dorans, N.J.; Flaugher, R.; Green, B.F.; Mislevy, R.J. *Computerized adaptive testing: A primer*; Routledge, 2000.
 96. Woolf, B.P. *Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning*; Morgan Kaufmann, 2010.
 97. Friedman, L.; Ahuja, S.; Allen, D.; Tan, T.; Sidahmed, H.; Long, C.; Xie, J.; Schubiner, G.; Patel, A.; Lara, H.; et al. Leveraging Large Language Models in Conversational Recommender Systems. *arXiv preprint arXiv:2305.07961* **2023**.
 98. Patel, N.; Nagpal, P.; Shah, T.; Sharma, A.; Malvi, S.; Lomas, D. Improving mathematics assessment readability: Do large language models help? *Journal of Computer Assisted Learning* **2023**, 39, 804–822. <https://doi.org/https://doi.org/10.1111/jcal.12776>.
 99. Zhang, Y.; Ding, H.; Shui, Z.; Ma, Y.; Zou, J.; Deoras, A.; Wang, H. Language models as recommender systems: Evaluations and limitations. In Proceedings of the NeurIPS 2021 Workshop on I (Still) Can't Believe It's Not Better, 2021.
 100. Lim, L.A.; Dawson, S.; Gašević, D.; Joksimović, S.; Pardo, A.; Fudge, A.; Gentili, S. Students' perceptions of, and emotional responses to, personalised learning analytics-based feedback: an exploratory study of four courses. *Assessment & Evaluation in Higher Education* **2021**, 46, 339–359. Publisher: Routledge _eprint: <https://doi.org/10.1080/02602938.2020.1782831>, <https://doi.org/10.1080/02602938.2020.1782831>.
 101. Bonner, E.; Lege, R.; Frazier, E. Large Language Model-Based Artificial Intelligence in the Language Classroom: Practical Ideas For Teaching. *The Journal of Teaching English with Technology* **2023**, 2023. <https://doi.org/10.56297/BKAM1691/WIEO1749>.
 102. DiCerbo, K. Building AI Applications Based on Learning Research [Webinar]. <https://www.youtube.com/watch?v=ugyfdjI9NEk>, 2023. Accessed: 07.14.2023.
 103. Kasneci, E.; Seßler, K.; Küchemann, S.; Bannert, M.; Dementieva, D.; Fischer, F.; Gasser, U.; Groh, G.; Günnemann, S.; Hüllermeier, E.; et al. ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences* **2023**, 103, 102274.
 104. Tlili, A.; Shehata, B.; Adarkwah, M.A.; Bozkurt, A.; Hickey, D.T.; Huang, R.; Agyemang, B. What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments* **2023**, 10, 15.
 105. Mathrani, A.; Susnjak, T.; Ramaswami, G.; Barczak, A. Perspectives on the challenges of generalizability, transparency and ethics in predictive learning analytics. *Computers and Education Open* **2021**, 2, 100060.
 106. Barros, T.M.; Souza Neto, P.A.; Silva, I.; Guedes, L.A. Predictive models for imbalanced data: A school dropout perspective. *Education Sciences* **2019**, 9, 275.
 107. Yan, L.; Sha, L.; Zhao, L.; Li, Y.; Martinez-Maldonado, R.; Chen, G.; Li, X.; Jin, Y.; Gašević, D. Practical and Ethical Challenges of Large Language Models in Education: A Systematic Literature Review, 2023, [arXiv:cs.CL/2303.13379].
 108. Truong, T.L.; Le, H.L.; Le-Dang, T.P. Sentiment analysis implementing BERT-based pre-trained language model for Vietnamese. In Proceedings of the 2020 7th NAFOSTED Conference on Information and Computer Science (NICS). IEEE, 2020, pp. 362–367. <https://doi.org/https://doi.org/10.1109/NICS51282.2020.9335912>.
 109. Khosravi, H.; Shum, S.B.; Chen, G.; Conati, C.; Tsai, Y.S.; Kay, J.; Knight, S.; Martinez-Maldonado, R.; Sadiq, S.; Gašević, D. Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence* **2022**, 3, 100074. <https://doi.org/https://doi.org/10.1016/j.caeai.2022.100074>.

110. Knight, S.; Shibani, A.; Abel, S.; Gibson, A.; Ryan, P.; Sutton, N.; Wight, R.; Lucas, C.; Sandor, A.; Kitto, K.; et al. AcaWriter: A learning analytics tool for formative feedback on academic writing. *Journal of Writing Research* **2020**, *12*, 141—186. <https://doi.org/https://doi.org/10.17239/jowr-2020.12.01.06>.
111. Kochmar, E.; Vu, D.D.; Belfer, R.; Gupta, V.; Serban, I.V.; Pineau, J. Automated personalized feedback improves learning gains in an intelligent tutoring system. In Proceedings of the Artificial Intelligence in Education: 21st International Conference, AIED 2020, Ifrane, Morocco, July 6–10, 2020, Proceedings, Part II 21. Springer, 2020, pp. 140–146.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.