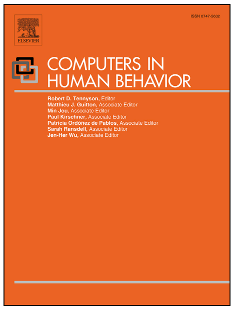
# Journal Pre-proof




The End is the Beginning is the End: The closed-loop learning analytics framework

Michael Sailer, Manuel Ninaus, Stefan E. Huber, Elisabeth Bauer, Samuel Greiff

PII: S0747-5632(24)00173-0

DOI: https://doi.org/10.1016/j.chb.2024.108305

Reference: CHB 108305

To appear in: *Computers in Human Behavior*

Received Date: 2 December 2023

Revised Date: 29 March 2024

Accepted Date: 16 May 2024

Please cite this article as: Sailer M., Ninaus M., Huber S.E., Bauer E. & Greiff S., The End is the Beginning is the End: The closed-loop learning analytics framework, *Computers in Human Behavior*, https://doi.org/10.1016/j.chb.2024.108305.

# The End is the Beginning is the End: The closed-loop learning analytics framework

*Michael Sailer[1*], Manuel Ninaus[2,3*], Stefan E. Huber[2*], Elisabeth Bauer[4*], & Samuel Greiff[5]*

[1]Learning Analytics and Educational Data Mining, University of Augsburg, Augsburg, Germany

[2]Department of Psychology, University of Graz, Graz, Austria

[3]LEAD Graduate School & Research Network, University of Tübingen, Tübingen, Germany

[4]TUM School of Social Sciences and Technology, Technical University of Munich, Munich, Germany

[5]Department of Behavioural and Cognitive Sciences, University of Luxembourg, Luxembourg

*These authors contributed equally to the manuscript.



## Author Note

The authors have no conflicts of interest to declare.

Correspondence concerning this article should be addressed to Manuel Ninaus, Universitätsplatz 2, 8010 Graz, Austria, manuel.ninaus@uni-graz.at

## Credit Author Statement

**Michael Sailer**: Conceptualization, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing

**Manuel Ninaus**: Conceptualization, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing

**Stefan Huber**: Investigation, Methodology, Writing – original draft, Writing – review & editing

**Elisabeth Bauer**: Investigation, Methodology, Writing – original draft, Writing – review & editing

**Samuel Greiff**: Validation, Writing – review & editing



## Declaration of generative AI in scientific writing

During the preparation of this work the authors used ChatGPT 4.0 for suggestions on wording and rephrasing. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

## Abstract

This article provides a comprehensive review of current practices and methodologies within the field of learning analytics, structured around a dedicated closed-loop framework. This framework effectively integrates various aspects of learning analytics into a cohesive framework, emphasizing the interplay between data collection, processing and analysis, as well as adaptivity and personalization, all connected by the learners involved and underpinned by educational and psychological theory. In reviewing each step of the closed loop, the article delves into the advancements in data collection, exploring how technological progress has expanded data collection methods, particularly focusing on the potential of multimodal data acquisition and how theory can inform this step. The processing and analysis step is thoroughly reviewed, highlighting a range of methods including machine learning and AI, and discussing the critical balance between prediction accuracy and interpretability. The adaptivity and personalization step examines the current state of research, underscoring significant gaps and the necessity for theory-informed, personalized learning interventions. Overall, the article underscores the importance of interdisciplinarity in learning analytics, advocating for the integration of insights from various fields to address challenges such as ethical data usage and the creation of quality learning experiences. This framework and review aim to guide future research and practice in learning analytics, promoting the development of effective, learner-centric educational environments driven by balancing data-driven insights and theoretical understanding.


*Keywords:* learning analytics, multimodal, artificial intelligence, education, adaptivity, personalization

## 1. The End is the Beginning is the End: The Closed-Loop Learning Analytics Framework

Learning analytics (LA) refers to the “measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs” (Long & Siemens, 2011). To achieve these aims, the field of LA applies methods and theories from data science, education, psychology, and instructional design (Gašević et al., 2015). In this article, we will introduce a framework that integrates different perspectives and foci in LA research, aiming to conceptualize new directions and to provide a holistic view of this interdisciplinary field of research.

The genesis of LA was marked by excitement over massive open online courses (MOOCs), with initial research heavily focused on data-driven approaches to pattern detection in learners' log data, mainly to predict learning outcomes. However, this approach, while beneficial in specific contexts (Alonso-Fernández et al., 2019), faced challenges in generalization and raised ethical and privacy concerns (Drachsler, 2018). The field's early enthusiasm has since matured into a more balanced view, recognizing the need for a nuanced approach that harmonizes technological advancements with ethical and theoretical considerations.

Currently, LA is navigating a landscape marked by technological innovations, such as wearables, advanced machine learning techniques (e.g., Ninaus & Sailer, 2022; Ouhaichi et al., 2023), and large language models (e.g., Kasneci et al., 2023), as well as a renewed emphasis on educational and psychological theory that has traditionally rather been neglected in the field of LA and yet builds the cornerstone of learning in the fields of education and psychology. This shift signifies a movement from a predominantly data-driven focus to a more holistic approach that values theoretical grounding and the thoughtful design of learning environments.

Our proposed closed-loop framework reflects these contemporary developments. The loop encapsulates the journey from data collection, processing and analysis to the design of adaptive and personalized learning environments. All of these steps are interconnected by the learners, who are the source of data collection in adaptive and personalized learning environments. Similar closed-loop systems are known from other feedback-rich systems such as bio-/neurofeedback (e.g., Ninaus et al., 2013), brain-computer interfaces (e.g., Kober et al., 2018), and AI-supported learning systems (Ninaus & Sailer, 2022). Furthermore, a similar loop for learning analytics has also already been discussed by Clow (2012), which has recently been used to discuss the role of generative AI in LA (Yan et al., 2024). The current Closed-Loop Learning Analytics Framework builds upon these systems/frameworks and, in addition, emphasizes the use of educational and psychological theory in the steps of applying LA as well as using data in deepening our understanding of learning processes (see also Wong et al., 2019 for educational theory development using iterative loops). Our Closed-Loop Learning Analytics Framework aims to bridge the gap between practical application and theoretical advancement, highlighting the symbiotic relationship between data-driven insights and educational as well as psychological theory in a largely iterative process. A visualization of the closed-loop system, which is inspired by Ninaus and Sailer (2022), is shown in Figure 1.

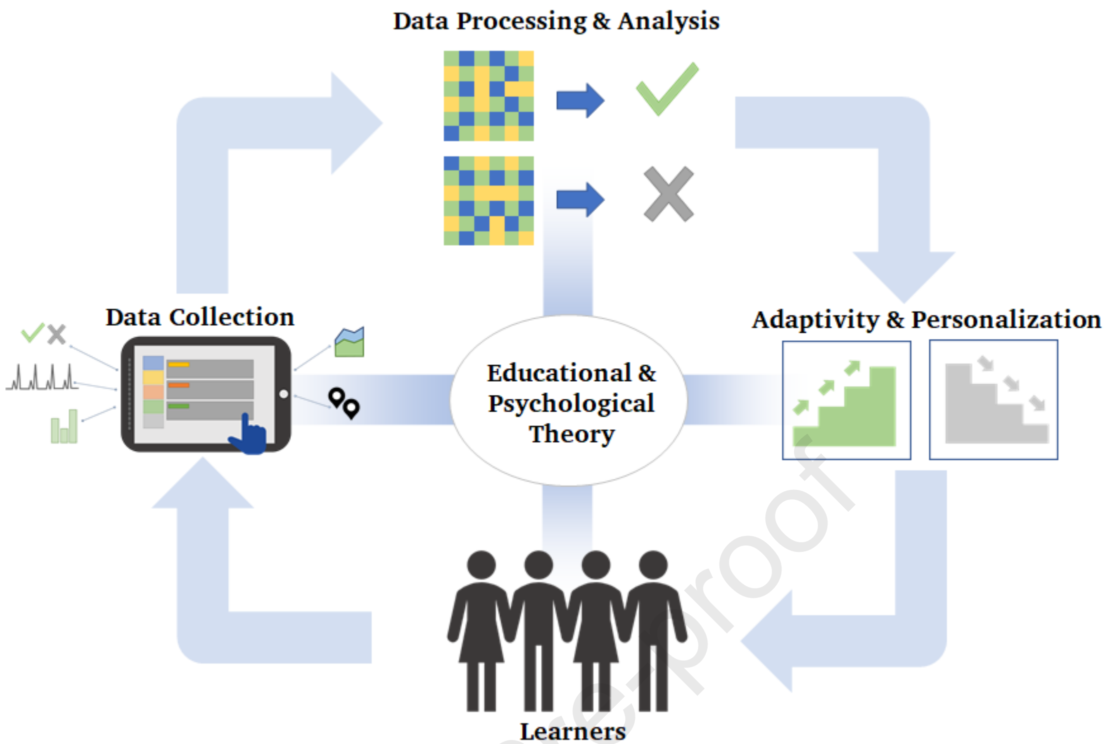



*Figure 1. Closed-Loop Learning Analytics Framework with the Steps (1) Data Collection, (2) Data Processing and Analysis, and (3) Adaptivity and Personalization That are Connected by the Learners as the Source of Data Collection and the Recipients of Personalized Learning Environments. All Components are Intertwined by Educational and Psychological Theory.*



In giving an overview of the existing literature and developing an overarching framework, we contribute to the ongoing discourse in LA, by offering a conceptualization that captures the essence of its interdisciplinary nature and addresses the current challenges and opportunities within the field.

## 2. The Closed-Loop Learning Analytics Framework and Review of the Literature

Learning is a complex phenomenon influenced by physiological processes, personal dispositions, social interactions, environmental aspects, and organizational constraints. Learning that takes place in digital environments leaves behind 'digital traces' (Giannakos & Cukurova,

2023) stemming from the interaction of the learner with the digital environment. Data relevant for the learning process, however, can also come from, for instance, wearable sensors, video cameras, audio recording devices, or from conventional questionnaires and tests (e.g., Di Mitri et al., 2018; Gray & Bergner, 2022; Nebel & Ninaus, 2019; Samuelsen et al., 2019). LA aims to utilize these data to make inferences about underlying processes, with the primary goal to better understand and ideally to support learning. In other words, LA aims to measure learning and related processes to better describe, understand, predict – and in an optimal scenario – optimize learning (Gray & Bergner, 2022). Therefore, the foundation of every LA system rests upon the collection of data, which constitutes the first step of our closed-loop framework (see 2.1). The second step involves processing and analyzing this data to extract meaningful insights. This stage delves into the methodologies employed in transforming raw data into actionable knowledge (see 2.2). The third step of our closed-loop framework revolves around the application of these insights. It explores how adaptivity and personalization, grounded in data-driven understanding, can enhance the learning experience, tailoring it to individual needs and preferences (see 2.3).

The closed loop is interconnected by the learners. Learners are the recipients of personalized learning experiences (see 2.3) as well as the sources for data collection (see 2.1), closing our proposed loop. As educational and psychological theory can inform, guide, and interconnect all steps described in the closed-loop system, it is set in the center of our loop.

In addition to introducing and specifying the closed-loop framework as displayed in Figure 1, we explore current trends and the state of research on LA, as guided by the *Closed-Loop Learning Analytics Framework* within a narrative review. Our review is informed by a diverse body of literature spanning psychology, education, and data science, selected to illustrate key themes and insights within each component of our framework. By providing this overview, we will focus on evidence from systematic reviews and meta-analyses that refer to different steps of our closed loop and supplement these by single empirical studies of particular relevance for the understanding of the closed-loop framework. Our review is structured to first present a broad spectrum of approaches within each section, followed by a focused discussion on the prevailing trends and innovations that are shaping the future of LA. This approach allows us to provide a comprehensive overview of the field, highlighting interdisciplinary contributions, and suggesting directions for future inquiry. While our literature search was not systematic as such, the examples chosen are representative of broader trends and debates within and across contributing disciplines. As our framework focusses on learning processes in digital learning environments, we also adopt this perspective for our narrative review.

**2.1. Data Collection**

This section delves into the diverse methodologies and tools employed in the collection of learning-related data. We begin by exploring a range of approaches, from traditional methods like questionnaires and tests to recently emerging techniques involving digital trace data and sensor-based collection. Subsequently, we shift our focus to current trends, particularly those that are redefining the efficiency, scope, and ethical considerations of data collection in educational settings.

Acquiring data during learning or in learning environments has become easier over the years due to technological developments in the field of wearable sensors, Internet of Things, and the general digitalization of learning. Importantly, though, the connection between the acquired data and the actual learning process and learning-related states, such as emotion or motivation, is not always straightforward. In fact, mostly we are interested in latent variables or – more specifically – learning constructs (e.g., Gray & Bergner, 2022), such as student attitudes (e.g., Leifheit et al., 2020; Suárez et al., 2019), affect (e.g., Cloude et al., 2022; Ninaus, Greipl, et al., 2019; Pekrun et al., 2017), or (changes in) abilities and skills (e.g., Tsarava et al., 2022). These variables cannot always be directly observed but need to be inferred from metrics that are in fact directly observable (i.e., from manifest indicators; e.g. Borsboom et al., 2003). The scores on a test or the ratings on a questionnaire, for instance, are directly observable and therefore might act as measurements for a specific variable or construct that is latent and is not directly observable. In the following paragraphs, we will provide a short overview of frequently used data sources and modalities used in LA research and address their specific advantages and disadvantages (for a more comprehensive account see, e.g., Samuelsen et al., 2019). In their entity, they relate to the step “Data Collection” in the close-loop framework (cf. Figure 1).

An overview of different data collection approaches sorted by their degree of obtrusiveness is summarized in Figure 2.

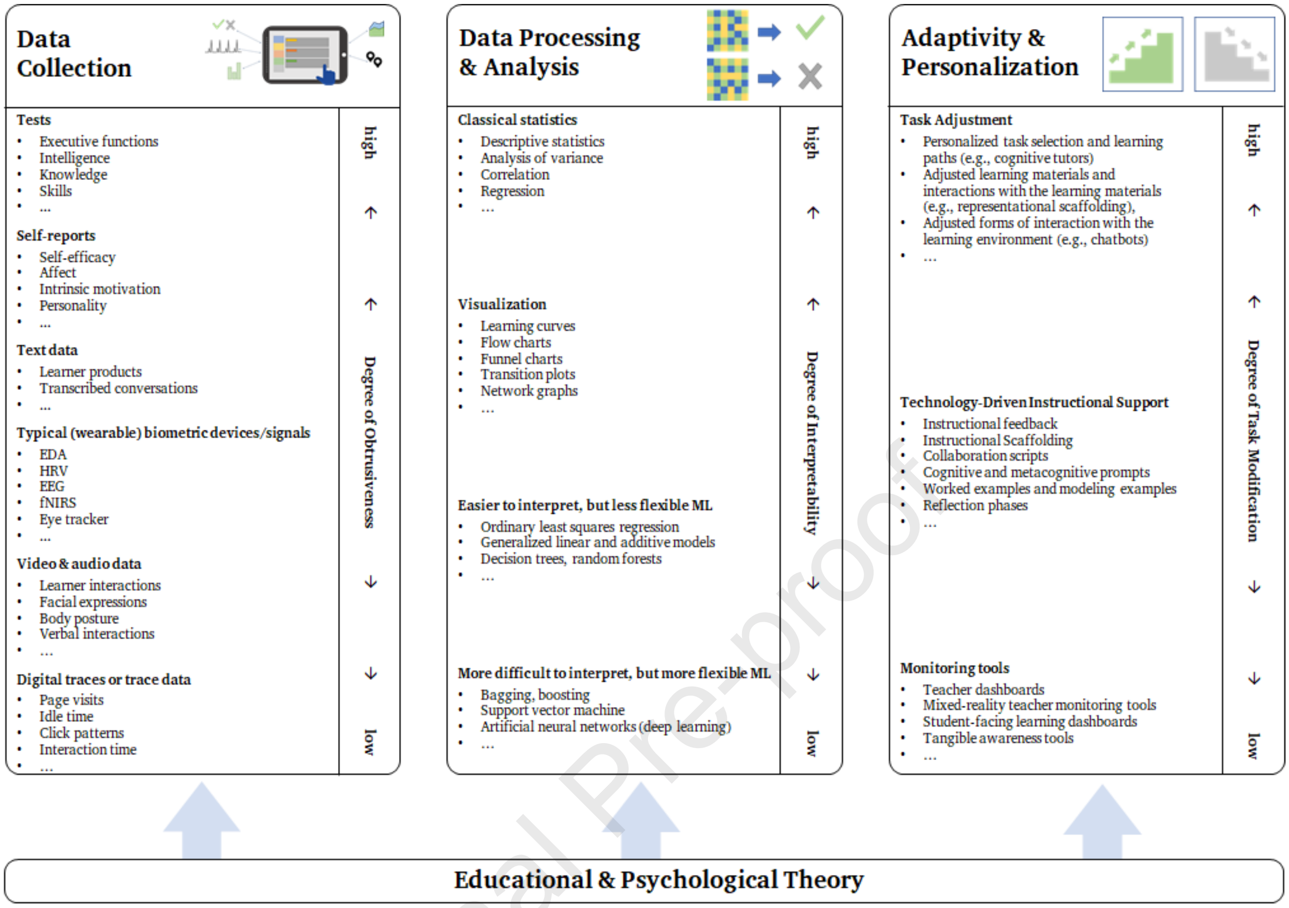



*Figure 2: Zoomed-in Section of the Closed-Loop Learning Analytics Framework Including an Exemplary Overview of Steps in the Framework Classified by Their Degree of Obtrusiveness (Data Collection), Transparency (Data Processing and Analysis), and Task Modification (Adaptivity and Personalization)*

### 2.1.1 Approaches for Data Collection in Learning Analytics Research

**Self-report questionnaires and tests** have the benefit of being reliable and valid – if they are psychometrically evaluated and have been carefully validated. That is, in the case of evaluated tests and questionnaires, psychometric properties, such as reliability and validity, are known beforehand and these instruments have been shown to serve the purpose they were developed for in many respects. They usually target a clearly defined latent construct with good theoretical grounding. However, self-report measures can be biased because of demand effects (e.g., Bernecker & Ninaus, 2021) or by inaccurate recall (Parry et al., 2021) and memory effects, such as recency (Freeman et al., 1999). Furthermore, self-report measures or other assessments that are somewhat separated from the learning process itself can disrupt the learner and thereby disturb the subjective experience in question and learning experience in general. Self-report questionnaires and tests can also only provide a rough estimate of the actual (dynamic; e.g., affective) processes involved in learning due to their lack of a usually low sampling frequency (Nebel & Ninaus, 2019). Further, research indicates that response tendencies can make it even more difficult to gather reliable information. Even as little as 5% of data, such as missing values and response biases, can markedly change the underlying structure of factors (Arias et al., 2023). Yet, reliable and valid questionnaires are vital for assessing rather stable learner variables, such as personality traits (Matcha, Gašević, et al., 2020).

For other data sources, such as **digital traces or trace data**, which are mostly automatically recorded from a learner when interacting with a digital learning environment, such as page visits, click patterns, or time spent in a learning management system – metrics often referred to as clickstream data – are usually higher in their sampling frequency and learners do not need to be probed (see also Shute & Ventura, 2013 for the related concept of stealth assessment). However, these data are not that easily linked to a specific (latent) construct and theoretical integration often becomes challenging or even impossible. For instance, while regular visits to the course material in an online course can be used to predict students' achievement performance (e.g., Li et al., 2020), this metric might not be a high-quality measure of the ability one is supposed to learn in the course. That is, one of the main challenges of using trace data is to draw clear inferences and provide explanations (for a more comprehensive account on this issue see Gray & Bergner, 2022). Further, trace data are often context specific, meaning, for instance, that one or several metrics in some course A of a learning management system are indicative for student engagement, whereas these metrics might not be indicative for engagement in some other course B. In this context, the specific behavioral components reflected in the trace data that constitute student engagement are contingent upon the unique structure or demands of each course (Motz et al., 2019). Consequently, in order to establish more valid metrics for a latent learning construct, theory-guided and deliberated decisions are necessary to guide the selection of metrics and learning constructs relevant for a given learning process. In the context of educational technologies, this entails deliberately designing data collection tools and, consequently, the trace data they acquire to closely conform to the learning process and the related learning construct(s) (Gray & Bergner, 2022), which is no trivial endeavor and thus barely done.

Data coming from **wearables or biometric devices** can provide information about the environment, such as temperature, location, humidity, but also rather personal information of human behavior (Di Mitri et al., 2018). While the use of wearables to monitor physical activity (e.g., smartwatches, fitness trackers, etc.; for a review see Gal et al., 2018) including the use of biometrics (e.g., heart rate monitoring) has become quite common in today's society, the use of such data to record and optimize learning activities is still relatively uncommon in educational contexts (Fortenbacher et al., 2019), particularly regarding real-time personalization of learning tasks (see also Section 2.3). Importantly though, physiological data from biometric devices can carry highly relevant information for learning and provide indicators for learning (e.g., Di Mitri et al., 2018; Nebel & Ninaus, 2019; Ochoa, 2022; Schneider et al., 2015) with high sampling rates allowing for capturing dynamic changes during the learning process. Such measures can be used, for instance, for dynamic difficulty adjustments as demonstrated by Ninaus, Tsarava et al. (2019), where learners' heart rate was used for adapting the difficulty in a game-based emergency simulation in real-time. In another study by Giannakos et al. (2020), physiological data captured from a wrist-worn wearable was successfully used to predict users learning experience.

The acquisition of **electrodermal activity (EDA) and heart rate variability (HRV)** is a common approach in LA research, for instance, to investigate self-regulated learner engagement (e.g., Wiedbusch, Dever, et al., 2023). For instance, a study by Gao et al. (2020) investigated the use of EDA metrics to capture cognitive, behavioral, and emotional engagement in learners. The results indicated that these engagement levels during class instruction could be detected with 79% accuracy using 12 EDA metrics along with other physiological measures. However, collecting EDA and HRV data poses challenges due to the intrusiveness of the instrumentation required. While devices like smartwatches offer unobtrusive data collection, more sophisticated instruments provide greater accuracy (e.g., Henrie et al., 2015). Considerations for environmental conditions (e.g., temperature in the room) and individual physiological and lifestyle differences further complicate data collection. In addition, the unobtrusiveness of data collection is also entangled with data privacy concerns and transparency requirements, which are discussed comprehensively by Drachsler (2018) as well as Drachsler & Greller (2016). Despite these challenges, EDA and HRV metrics offer fine-grained data on learning relevant states and high-quality sensors used to acquire these data have become more affordable over recent years.

Another relevant source of physiological data is the organ in which learning is actually happening – the human brain. The most common device to acquire **neurofunctional activity** is the electroencephalogram (EEG), which non-invasively records the electrical activity of the brain. Functional near-infrared spectroscopy (fNIRS) measures hemodynamic changes in the blood flow of the brain (Ferrari & Quaresima, 2012) and thus utilizes a slower brain signal as compared to EEG (for a review see Ninaus, Kober, Friedrich, Dunwell, et al., 2014). There are other devices to measure brain activity, such as functional magnetic resonance imaging or Magnetoencephalography, but these measures are hardly applicable for naturalistic learning scenarios or educational settings. Consequently, while providing crucial information about the learning process and underlying mechanism, such as the coupling of cognition and emotion in game-based learning (e.g., Greipl, Klein, et al., 2021), they are barely used in the field of LA research. EEG and fNIRS are more popular in the field of LA (e.g., Gao et al., 2020; Giannakos et al., 2019; Mills et al., 2017; Ninaus, Kober, Friedrich, Neuper, et al., 2014; Witte et al., 2015) as they are better suited to be used in naturalistic learning settings also due to recent advancement in wearable or portable neuroimaging devices (Zhan et al., 2023). For instance, Mills and colleagues (2017) used a portable 24-channel EEG headset to collect brain activity of learners while they were learning with an intelligent tutoring system to predict different levels of cognitive load. Nevertheless, bringing neuroscientific devices into naturalistic learning settings, for instance, the classroom, comes with many challenges such as preparation time for EEG caps or NIRS probe sets, artifacts produced by movements or other electrical devices (for a review see Janssen et al., 2021).

Data from **eye tracking** devices are another important factor in LA research and are easier to acquire than using neuroimaging devices. Metrics derived from professional eye trackers, such as micro-saccades, fixations, or pupil dilation can be indicative, for instance, of learners cognitive load when playing a serious game (e.g., Appel et al., 2021) or changes in gaze direction can provide helpful measures for attention patterns when learning from videos (Srivastava et al., 2021). Importantly, recent studies demonstrated that for certain scenarios webcam-based eye tracking might be sufficient to detect learning constructs relevant to the learning process (Hutt et al., 2023; Khosravi et al., 2022). For instance, Hutt and colleagues (2023) could predict mind wandering using gaze measurements derived from a conventional webcam during an online reading comprehension task.

**Video data** of learners can provide us with, among other things, information on learners affective states by detecting facial expressions (Cloude et al., 2022; Frenzel et al., 2024; Ninaus, Greipl, et al., 2019) or other bodily expressions (Greipl, Bernecker, et al., 2021; Riemer et al., 2017). Video or image data can also be used to capture data from the context of where or under which circumstances learning occurs (e.g., presence/absence of other learners and their interaction). For instance, Vuorenmaa and colleagues (2023) used videotaped sessions of secondary school students while performing collaborative physics tasks to investigate social interaction processes beneficial for learning in small groups.

**Text data**, for instance, coming from text-based (or transcribed) conversations with a virtual agent or other learners, are another important data source. Large quantities of text-based data can be analyzed using text mining and natural language processing. Which can provide very basic descriptive information, such as the number of words, words per sentences, but also more complex information on semantic content or sentiment of a text (Allen et al., 2022; Dowell & Kovanović, 2022). The latter, was, for instance, used in detecting affective states of learners in MOOCs (Chaplot et al., 2015; Dalipi et al., 2021). Moreover, text data can also constitute the learning product of a learner, such as scientific argumentation, and thereby provide a measure for learning outcomes (e.g., Bauer et al., 2022). With the rise of Large Language Models (LLMs) and artificial intelligence (AI) in education (Kasneci et al., 2023) new use cases will emerge improving eventually on current methodologies (e.g., automatic transcription of video/audio data; see Yan et al., 2024).

#### 2.1.2 Current Trends - Harnessing New Technologies and Multimodal Approaches

Each of the mentioned data sources and metrics have their own advantages and disadvantages and should be carefully selected keeping the goal of the specific investigation in mind. In an optimal case, selection should be based on educational and psychological theory to increase the overlap between the actual measures and the construct of interest, usually learning in one way or the other. However, this is not always feasible, and measures are selected based on availability and legitimate practical considerations such as time and monetary constraints.

Importantly, single metrics or single data sources usually provide only a partial and noisy picture of learners' actions and the underlying (learning) processes. This is further exacerbated by incomplete or missing data (e.g., due to failing sensors or other technical difficulties). Having data available from multiple sources can partly overcome these issues (e.g., Bosch, 2015; Di Mitri et al., 2018), because the signal-to-noise ratio can be improved when different metrics that tap into the same construct, at least to some extent, are employed (i.e., multimodal measurement of the same construct). As such, joy or frustration during learning may be inferred from facial expressions captured with a webcam (e.g., Cloude et al., 2022; Ninaus, Greipl, et al., 2019) but also from self-report measures (e.g., Bosch, 2015; Wortha et al., 2019).

Further, when relying on a single type of source or even a single metric, the complex, multidimensional process of learning is oversimplified and data may be misinterpreted because of, among other things, missing contextual information such as where and under which circumstances learning occurred (Ochoa, 2022; Selwyn, 2019; Wiedbusch, Dever, et al., 2023). To capture the multidimensional nature of learning, modern LA approaches make use of data from different sources to obtain a more holistic understanding of learning processes (Molenaar et al., 2023; Spikol et al., 2017). These “multimodal data” have become more prominent in general and in LA field in particular (see multimodal learning analytics; Blikstein, 2013). Crucially, different data sources can provide unique and complementary perspectives on learning constructs. A combination of sources and metrics is commonly referred to as “fusion” or integration of such data and can increase the probability to measure the latent construct of interest or different aspects thereof.

For instance, Giannakos and colleagues (2019) demonstrated that click-stream data alone are predictive of learning performance. However, integrating various other data sources, such as eye-tracking, EEG, videos of learners' faces, and wristband data (e.g., heart rate, EDA), could significantly improve learning performance prediction. Importantly, data from different sources might also be interpreted in different directions when considered separately, but only the combination of multiple data sources might allow for clear conclusions regarding a latent construct that tells us something about learning and its underlying process. As such, Taub et al. (2017) found that analyzing data from multiple sources together rather than analyzing data sources separately produced different results in a game-based learning scenario. This demonstrates the complementary role different data sources can have to better understand the learning process and predict learning outcomes (see also Section 2.2). Further, combining multiple data sources seems to be particularly relevant when measuring multidimensional constructs, such as flow or engagement, because such constructs can influence learning in various ways (e.g., Papamitsiou et al., 2020; Wiedbusch, Dever, et al., 2023).

Unimodal traces can sometimes be enough to measure the desired learning construct. Most often, to increase accuracy and prediction rates, multiple unimodal traces need to be fused together (e.g., Ochoa, 2022). However, the fusion of data from various sources is no trivial task. For instance, in collaborative learning scenarios gaze direction from individuals needs to be fused together in order to be able to detect when two or more learners at the same time intersect inside a given area of interest (see Ochoa, 2022). Furthermore, if EEG data are supposed to be fused together with heart rate, EDA or fNIRS, it needs to be considered that these different signals have different latencies. For instance, fNIRS – a signal relying on hemodynamic changes in the brain – is substantially slower in showing a reaction to a stimulus (hemodynamic response peaks at around 4-6 seconds after a given stimuli; Huettel et al., 2009) than EEG, which responds more or less immediate to a stimulus because the electrical brain activity is being measured.

Furthermore, the collection of data from multiple sources also needs to consider the varying sampling frequencies and thus the temporal alignment of multimodal data. For instance, suggested sampling frequencies for EDA (i.e., 200-400 Hz; Braithwaite et al., 2013) or HRV (i.e., 500 Hz; Berntson et al., 1997), which vary however due to their respective use-case (e.g., Bent & Dunn, 2021), are vastly different to conventional video data or even self-reports, which provide one data point every other minute or – more likely – every half an hour or hour. In this context, another crucial temporal aspect has to be considered – the temporal granularity difference between metrics of the same data source (Azevedo & Gašević, 2019). For instance, comparing more fine-grained (e.g., fixation durations) with aggregated data of the same source (e.g., heatmaps) strongly influences which learning construct can or should be investigated (e.g., shifts of attention in a single moment vs. overall attentional investment into a learning task).

Effectively, the selection of data sources and metrics should be informed by educational and psychological theory with the aim to obtain a level of accuracy that is acceptable for the purpose at hand. This, of course, differs substantially, for instance when the intention is to provide individual feedback (high accuracy required) or when to investigate group differences (low accuracy might be acceptable). This implies, for instance, that when one is interested in student dropout across semesters compared to attention processes in a specific learning task, different data sources and metrics are employed. That is, comparing or fusing together different data sources comes with different challenges but can be supported by theory-guided decisions as to which data sources should be acquired in the first place avoiding data-fishing approaches. For instance, cognitive load theory, which is a theory of human cognitive architecture to design instructional procedures (Sweller, 2024) and related research (e.g., Ayres et al., 2021; O. Chen et al., 2023; Vanneste et al., 2021) offer manifold and helpful discussions on how to validly collect data on facets of learners' cognitive load and how to combine those measures of data collection.

Further, a theory-guided selection of data to be acquired also facilitates the next step in our Closed-Loop Learning Analytics Framework - data processing and analysis.

### 2.2 Data Processing and Analysis

In this section, we examine the critical phase of processing and analyzing the collected data with relation to the source it derives from and the measures employed. After an overview of various data processing techniques and analytical strategies, we highlight how advancements in areas such as machine learning and big data analytics are transforming the way educational data is interpreted and utilized.

A thorough understanding of the collected data and their relation to underlying learning processes can be considered a necessary prerequisite for successful adaptiveness of digital learning environments enabling optimal, personalized learning experiences (Ehlenz et al., 2022). In this context, data processing and analysis comprise all steps required to identify patterns in the recorded data that are relevant for the learning process. Data processing and analysis thus provide the methodological link between the underlying, collected data (see Section 2.1) and their successful utilization in adaptive learning systems (see Section 2.3).

Manual processing of possibly highly granular, temporally dense data from numerous sources is at least difficult, in certain cases, impossible at all. Therefore, LA makes extensive use of computational data science approaches employing frequently machine learning or statistical learning methods (Aldowah et al., 2019; Alonso-Fernández et al., 2019; Baker & Siemens, 2022; Banihashem et al., 2023; Bond et al., 2023; X. Chen et al., 2020; Du et al., 2021; Hilbert et al., 2021; Namoun & Alshanqiti, 2020; Romero & Ventura, 2020). Thus, large methodological overlaps exist with other research communities such as artificial intelligence in education (AIED), educational data mining (EDM), or user modeling, adaptation and personalization (UMAP) (Giannakos & Cukurova, 2023).

A few examples may illustrate the heterogeneity of data and methodological approaches regarding data processing and analysis in the framework of LA.

1. In a study by Appel et al. (2021), multiple metrics of eye tracking data (i.e., pupil diameter, blinks, fixations, microsaccades) were utilized in conjunction with random forest classification to train participant-specific classifiers for distinguishing between

low and high cognitive load using data from an N-back task. The weighted predictions from these classifiers were then applied to different participants playing a simulation game. This cross-task and cross-participant classification algorithm successfully predicted whether participants were playing a more or less difficult level of the simulation game.

2. Recently, Martin et al. (2023) used computational grounded theory (Nelson, 2020) for evaluating written student argumentations about the plausibility of competing chemical reactions. Using pre-trained LLMs in conjunction with deep neural networks, the authors obtained a machine-human score agreement with an accuracy of 87%.
3. Multimodal video analysis via convolutional neural network classification was used by Ocak et al. (2023) to analyze children's computational thinking as captured in visible modes of interaction. The high agreement between human and machine assessment obtained by the authors indicated that the computational approach could act as an additional quality assurance measure in such analyses.
4. Ouyang et al. (2023) employed hidden Markov modeling in conjunction with lag sequential analysis and frequent sequence mining to analyze collaborative knowledge construction (CKC) on the basis of audio recordings of group discussions. The results could reveal insights into both the multilevel characteristics and the dynamics of CKC. The authors concluded that the computational approach used to classify learner behavior into different CKC states could enable future automated learning systems to provide personalized scaffolding opportunities based on the detected patterns.

**2.2.1 Approaches for Data Processing and Analysis in Learning Analytics Research**

The given examples illustrate that many computational approaches exist in the framework of LA and they are difficult to subsume under a common set of categories. In fact, to structure the multitude of methods, different taxonomies have been suggested. A broad but widely-accepted one in statistical learning distinguishes between supervised and unsupervised algorithms (James et al., 2021).

Supervised algorithms aim to establish a functional relationship between a set of input variables (also known as covariates, features, or predictors) and a corresponding set of output variables (also known as target, response, or dependent variables) (Hilbert et al., 2021). The goal of establishing that functional relationship can be to accurately predict responses for future observations of inputs (prediction) or to better understand the relationship between the response and the predictors (inference) (James et al., 2021). The first three examples above (Appel et al., 2021; Martin et al., 2023; Ocak et al., 2023) employed supervised algorithms. Generally, supervised algorithms comprise a wide range of methods including classical statistical methods like linear or logistic regression, regularized regression such as LASSO or ridge regression, support vector machines, decision trees and random forests, or artificial neural networks (Hilbert et al., 2021). For a more comprehensive overview, see, for instance, James et al. (2021).

Supervised algorithms can be further decomposed into classifiers, regressors, and latent knowledge estimation (Baker & Siemens, 2022). Classifiers aim for categorizing each observation into one of several given, distinct categories. That means the output variable is qualitative (or categorical). A random forest (e.g., Appel et al., 2021) is an example of a classifier. For regressors, the target variable is quantitative (or continuous; i.e., it can be modeled by a real number). Linear regression is an example of a regressor. Latent knowledge estimation is a special subtype of classifiers (Baker & Siemens, 2022).

In contrast to supervised algorithms, unsupervised algorithms deal with data in which no output variables are available for a given set of input variables. The goal of unsupervised algorithms is to identify relationships or structures among input variables. By aiming for identifying distinct multilevel and dynamic patterns of CKC in audio data (Ouyang et al., 2023), the fourth example given above belonged to this category. Unsupervised algorithms are also known as structure discovery algorithms. Typical approaches in LA belonging to this class of algorithms include clustering, latent profile or latent class analysis, correlational or factor analysis, domain structure discovery, and network analysis (Baker & Siemens, 2022). These methods can detect patterns of typical learner characteristics and learning process characteristics, with studies using, for example, latent profile analysis (e.g., Nickl et al., 2022; Radkowitsch et al., 2023) or epistemic network analysis (e.g., Omarchevska et al., 2022; Shaffer, 2017).

Variations of structure discovery methods with the goal to discover relationships between variables in very large datasets are known as relationship mining and represent historically a core category of EDM research (Baker & Yacef, 2009). According to Baker and Siemens (2022), relationship mining comprises typically four types: association rule mining (i.e., a method aiming to discover interesting relations between variables in large data sets), sequential pattern mining (i.e., a method specialized in identifying patterns in sequential data), correlation mining (i.e., a method for clustering relationships between variables rather than observations), and causal data mining (i.e., a method employing causal machine learning algorithms aiming for the identification of causal networks in data).

Between supervised and unsupervised algorithms, semi-supervised algorithms deal with data in which only a fraction of input variables is associated with corresponding output variables whereas the remaining fraction is not (James et al., 2021). Also situated in between supervised and unsupervised algorithms is another machine learning paradigm known as reinforcement learning dealing with sequential decision-making problems under the constraint of limited feedback (Van Otterlo & Wiering, 2012).

Besides machine learning algorithms, also classical statistical methods (e.g., descriptive statistics, analysis of variance) are frequently employed in LA (Du et al., 2021). The distinction between classical statistical methods and machine learning algorithms is, however, not clear-cut. For instance, ordinary-least-squares regression is usually subsumed under supervised algorithms whereas it also represents a classical statistical method. The same holds for correlational analyses.

Another methodology in LA, however, outside the scope of supervised and unsupervised algorithms, is discovery with models, which – in its simplest variant – involves two main phases (Hershkovitz et al., 2013). In a first phase, a model of some psychological construct or latent variable (e.g., boredom, confusion or self-regulative behavior) is obtained, typically using some prediction method (Baker & Siemens, 2022). In a second phase, this model is then used as a component in another analysis. For example, it is investigated which elements in a learning system are associated most with the development of boredom, whereas boredom is detected based on the model previously obtained. As outlined by Baker and Siemens (2022), discovery with models might also involve hierarchical structures of models, i.e., a model could be composed of other models which in turn could be composed of other models and so on.

Another methodology commonly employed in LA is distillation of data for human judgment, also known as visualization (Baker & Siemens, 2022). The focus of visualization is on comprehensibility of patterns associated with the learning process. Visualization comprises, for instance, learning curves (i.e., depicting performance as a function of time; Peddycord-Liu et al.,

2018) or illustrations of the classroom layout (Holstein et al., 2017). For instance, visualizing student grading history over school years was used to identify common patterns among successful and unsuccessful students (Bowers, 2010). The identified patterns were discussed to be utilized for providing timely support for students at risk of dropping out of school.

To summarize, data analysis in the framework of LA can draw on specialized methods from a broad arsenal of possibilities. Which of the various methods is used, depends, however, also on the goal of a study. If a study's goal is prediction, then supervised algorithms are the dominant option (Namoun & Alshanqiti, 2020).

To some extent, the exact choice of methods, however, also depends on contextual factors or the topic of investigation. Whereas in a systematic review generally aiming for the prediction of student performance (Namoun & Alshanqiti, 2020) support vector machines were found to be employed only in 2 of 62 studies, the same method turned out as the matter of choice in a review specifically focusing on the use of LA in programming courses (Omer et al., 2023). Generally, machine learning algorithms (support vector machines, K-nearest neighbor, naïve Bayes, decision trees and random forest, deep learning) appeared more dominant compared to classical statistical analyses (correlation, regression) in the latter (Omer et al., 2023) than in Namoun and Alshanqiti's systematic review (2020).

While the prediction of performance has been the most prolific theme in LA research, other upcoming directions comprise the identification and monitoring of students' learning progress, and utilizing analytic results for providing feedback or modeling of the learning process (Du et al., 2021). Given this diversity in epistemic goals, it is unsurprising that basically the full spectrum of computational approaches outlined above has been utilized in LA research. According to the review provided by Du et al. (2021), most studies were either of a descriptive or a predictive kind. Whereas classical statistical methods (45%), visualization (24%) and clustering (15%) were the most popular methods in descriptive studies, regression (79%) and decision trees (14%) were most commonly employed in the predictive ones (Du et al., 2021). The exact numbers may fluctuate, depending also on the specific focus of the considered review (Tepgec & Ifenthaler, 2022). The most prominent classes of data analytical methods in other, partially more specialized reviews (Aldowah et al., 2019; Alonso-Fernández et al., 2019; Tepgec & Ifenthaler, 2022; Viberg et al., 2018), however, comprise also a mixture of regression and classification approaches for supervised algorithms, and clustering, correlational and other classical statistical analyses for unsupervised algorithms, besides various visualization techniques. To a somewhat lesser extent, also the use of discovery with models, different relationship mining approaches, and various implementations of reinforcement learning algorithms are noted (Aldowah et al., 2019; Tepgec & Ifenthaler, 2022; Viberg et al., 2018).

**2.2.2 Current Trends - Embracing Diversity in Methods and AI in Learning Analytics**

Some trends seem to be emerging over the last decade. Banihashem et al. (2023) note that most studies used multiple data analytic approaches at various levels of their investigation. For instance, supervised algorithms were used for prediction, whereas visualization techniques were used at a descriptive level. Furthermore, combining advanced cluster analysis with different forms of sequence analysis was utilized for detection of cognitive and metacognitive processes from log data (Azevedo & Gašević, 2019). More generally, clustering and using both supervised and unsupervised algorithms has been noted as “an interesting way to drive our understanding forward of how to extract meaningful information from log data on learners’ strategic actions and the temporal development thereof” (Molenaar et al., 2023, p. 5). Adding in analysis of video data might give further insight or support automatization with advancements in video AI analytics (Duan et al., 2020).

This is in line with statistical authorities advocating the use of a plurality of methods for various reasons. As Wilcox (2022), points out, no method is always the best. Which method performs well depends on the situation. Specifically, on the underlying educational and psychological theory, type of collected data (see Section 2.1), as well as goals for optimizing learning (see Section 2.3). Furthermore, different methods can give different perspectives on the underlying data, and if used and interpreted carefully together, their combination may lead to deeper and better understanding of the processes generating those data.

James et al. (2021) note that there is always a trade-off between prediction accuracy and model interpretability. If, under some circumstances, prediction accuracy might be the entire purpose of using an algorithm, the choice might naturally fall towards a highly flexible approach like deep learning. However, the recent trend towards more interpretable and understandable methods (Hilbert et al., 2021) might suggest at least the accompanying use of more traditional, less flexible approaches like regression or complementing deep learning by techniques like network dissection (Zhou et al., 2019) allowing the human user some insight into the inner workings of the computational model. It should be noted though, that (full) transparency to human insight may not be a prerequisite for AI systems to be able to promote human understanding (Krenn et al., 2022). Ninaus and Sailer (2022) pointed out, interpretability and understandability are especially relevant for LA, because in that case, they are primary research goals. For example, consider LA researchers aiming for a better understanding which individual aspects of some intervention positively or negatively affect learning. To some extent, this underlies also the emphasis put on the use of causal models (Kitto et al., 2023) to bridge the gap between mere data processing and educational theory (Giannakos & Cukurova, 2023).

Even if prediction per se is the goal, the risk of overfitting – especially in the case of the most flexible approaches (James et al., 2021) – might advocate rather the use and comparison of various approaches varying in flexibility. Wagenmakers et al. (2021) also suggest the assessment of various plausible alternative statistical methods to gauge the extent to which a statistical conclusion is either fragile or sturdy, i.e., to assess the robustness of a certain result.

The trend towards multiple quantitative method approaches may be complemented further by an increase in mixed method studies, i.e., the combination of quantitative and qualitative approaches in one study. According to the review on LA in higher education provided by Viberg et al. (2018), only 17% of studies used a mixed method approach in 2012, the portion of studies doing so stabilized at about 30% in the following years. In a more recent review provided by Zhu et al. (2022), focusing, however, specifically on LA studies in the framework of MOOCs, a notable fraction of 37% of studies employing mixed methods were obtained. As pointed out already in one of this section's introductory examples (Martin et al., 2023), approaches such as computational grounded theory (Nelson, 2020) might leverage the potential of AI, and specifically machine learning and natural language processing, for scaling up qualitative research methods to the analysis of large-scale data (Tschisgale et al., 2023). This may in consequence open up perspectives for the inclusion of textual data (both in written and spoken form) in the step of pattern identification in LA.

Finally, complementing multi-method studies by including multiple input data streams opens the possibility of capturing the learning process in a comprehensive and holistic way. For instance, Giannakos et al. (2019) found that including additional data streams besides traditional log-data could reduce the error rate in predicting learning performance by more than 30%. Such encouraging findings provide a strong rationale to go beyond focusing merely on log data by including other data such as eye tracking, wearable cameras, gesture-recognition systems, facial expression analysis infrared-imaging and biosensors (Crescenzi-Lanna, 2020; Emerson et al., 2020) resulting in the emerging field of multimodal LA (MMLA). Informing LA by combining the information contained in two or more data streams holds the promise for a deeper understanding of the learning process (Sharma & Giannakos, 2020; Taub et al., 2017). The latter may form the basis for the development of digital, adaptive learning environments capable of detecting detrimental or beneficial learning behavior and of providing timely and appropriate support enabling individualized learning that builds on various sources of data and model input (Ehlenz et al., 2022; Emerson et al., 2020). By allowing data recording in a temporal and unobtrusive manner and facilitating dynamical analyses to model how multimodal processes may be intertwined, MMLA considerably extends earlier capabilities to interpret cognitive, affective and social processes underlying learning. By providing thus a both flexible and comprehensive methodological lens on learning, MMLA is seen to provide substantial potential to inform and extend on our current theoretical understanding of learning (Giannakos & Cukurova, 2023).

However, multi-method, multimodal studies also pose new challenges both for data recording and analysis (see also Section 2.1). One is the necessary alignment of different data streams with different degrees of granularity (Molenaar et al., 2023; see also Section 2.1.2). Synchronization and analysis of various data streams become even more challenging when it is the goal to study collaborative learning including data streams from several learners (Järvenoja et al., 2018). In such cases, further analytic approaches like recurrence quantification analysis, time series analysis, or hierarchical (linear) models have been utilized but there is unused potential that remains to be further elaborated upon (Cloude et al., 2020, 2022; Dindar et al., 2019; Wiedbusch, Lester, et al., 2023).

Another challenge is the integrated identification and utilization of patterns distributed over several data streams. One reported way of dealing with this issue (Molenaar et al., 2023) is the use of process mining techniques allowing to combine information from various data streams (Taub et al., 2021). Theoretically grounding identified event traces is preferable to mere data-driven approaches. For instance, in the study of Taub et al. (2021), a theory-driven approach, based on information processing theory (Winne, 2017) and the model of affective dynamics (D'Mello & Graesser, 2012), resulted in a detailed account of self-regulated learning processes that could not be obtained from frequency measures alone. Another example is the FLoRa project in which algorithms were validated by first training a multimodal data-driven process library on the basis of a theory-driven process library (Fan et al., 2022). As already outlined above to some extent, advances in natural language processing and generative AI are foreseen to facilitate especially so far typical steps of manual coding in the analysis of video, audio, and textual data (Molenaar et al., 2023). Besides supporting analyses of such, beforehand less clearly structured data, generative AI will extend illustrative, explanatory and interactive capabilities of LA (Yan et al., 2024). Further, generative AI will allow to extend the arsenal of automatically processed data beyond so far mainly utilized log and physiological data. The identified patterns within the combination of different data streams could then be forwarded to intelligent tutor systems such as MetaTutor (e.g., Azevedo et al., 2010). Moreover, retrieval-augmented generation (Lewis et al., 2020), could be utilized to link generative AI with complementary course material and learning theory to proactively support learners with additional resources or suggestions for learning strategies based on their current learning characteristics and performance (Yan et al., 2024). Hence, as outlined further by Yan et al. (2024), generative AI will especially open various new avenues regarding adaptivity and personalization to which we will turn next.

An overview of different data processing and analysis approaches sorted by their degree of interpretability is summarized in Figure 2.

### 2.3. Adaptivity and Personalization

Insights gleaned from data analysis can be applied to create adaptive and personalized learning experiences as outlined in Figure 1. We begin with revisiting different approaches to adaptivity and personalization, considering both theoretical underpinnings and practical implementations. Subsequently, we focus on trends and developments in data-driven and personalized teaching and learning.

#### 2.3.1 Approaches for Adaptivity and Personalization in Learning Analytics Research

Especially in its early years, the field of LA was sparked and driven by the massive interest in MOOCs and the large amounts of educational data generated therein (Gašević et al., 2015), resulting in a research focus on detecting patterns in educational data (e.g., to predict performance, engagement, or dropout rates; e.g., Leitner et al., 2017). However, this research was repeatedly criticized for neglecting educational and psychological theories and evidence from prior research and, consequently, characterized as offering only limited theoretical advancements as well as limited implications for practice given its often ad-hoc nature without close ties to theoretical consideration and specific use cases (e.g., Gašević et al., 2015; Nistor et al., 2015; Viberg et al., 2018). Currently, there is still a large portion of LA research that is neither focused on learning outcomes nor on learning interventions (Motz et al., 2023). Yet, more recent reviews suggest that research in the LA and related fields increasingly addresses matters of data-based personalized learning and adaptive instruction (Baker, 2023; Bond et al., 2023; Drugova et al., 2023; Du et al., 2021).

Personalized learning broadly refers to the data-based adjustment of any aspect of instructional practice to relevant characteristics of a specific learner, usually with the aim to provide optimal support to the learner (Bernacki et al., 2021; Tetzlaff et al., 2021). Personalizing learning requires (a) assessing the relevant learner characteristics and (b) deciding on potential interventions in light of the specific situation of the learner (Plass & Pawar, 2020; Tetzlaff et al., 2021). In an analog classroom, these tasks are usually considered the teacher's responsibility. However, also the students can take over agency for their learning by involving in self-assessment, internal feedback, and self-regulated learning (Andrade, 2019; Nicol, 2021; Zimmerman, 2002). In technology-enhanced learning environments, the assessment and the decision-making about interventions can be augmented or automated by means of data-based adaptivity. Adaptive educational technologies assess learner variables to accommodate users' specific needs with the goal of enhancing learning outcomes (Plass & Pawar, 2020). Reviews indicate that personalized learning that is supported by adaptive educational technologies can facilitate students' learning (Aleven, McLaughlin, et al., 2016; Bernacki et al., 2021; Ninaus & Nebel, 2021) – at least in the short-term, since research on long-term effects is widely missing, which is – given that the field has now existed for long over a decade – surprising and a relevant gap that needs to be closed rather urgently.

For the purpose of automated adaptation, educational technologies need to integrate recorded and preprocessed data (e.g., detected patterns) into a learner model, which denotes the technologies' assessment and representation of a learner – especially their learning prerequisites, learning processes, and predicted learning outcomes – to which the learning experience is adapted (Conati et al., 2018; Plass & Pawar, 2020). As described in the previous sections, there exist a multitude of data sources (see Section 2.1) and methods (see Section 2.2) that can be exploited to create learner models; however, not all are equally relevant for every personalized learning experience. To create meaningful learner models that allow adaptive educational technologies to purposefully address learners' needs, educational and psychological theories and evidence from prior research can be key to identify leverage points that offer high potential for advancing learning in a given context (e.g., Bauer et al., 2023; Van Der Graaf et al., 2023; Wiedbusch et al., 2021). Such leverage points can be found when focusing on why and in which regard learners struggle in a learning task. For example, in a learning scenario using peer feedback, learners' mindful processing of the feedback received by their peers might be limited through the complexity of the feedback message and low confidence in peers' competence (Berndt et al., 2018, 2022; Bolzer et al., 2015; Huisman et al., 2020; Patchan & Schunn, 2015), which may be addressed by providing targeted learner support (Alemdag & Yildirim, 2022; Bauer et al., 2023).

Although educational and psychological theories and research can offer ideas for specific learning scenarios, attempts to integrate such insights into systematic theoretical approaches of personalized learning still leave a lot of unresolved questions regarding on which basis to adapt which support in order to increase specific outcomes. One of the most systematic approaches consists of research on aptitude treatment interactions (ATI; Cronbach, 1957), that is, assessing interindividual differences in relevant learning prerequisites (aptitudes) to form groups of learners and testing whether interventions (treatments) yield different effects for these groups. This research resulted in insights, such as the expertise reversal effect (Kalyuga, 2007), which highlights that interventions that offer positive effects for novice learners can have negative effects on advanced learners. However, overall, ATI research is considered as not having achieved robust results that could be translated into generalizable approaches for personalized learning (Plass & Pawar, 2020; Tetzlaff et al., 2021; Tobias, 1989). A different approach can be found in the research on intelligent tutoring systems (ITS). These computer-based learning environments help students master knowledge and skills through step-by-step guidance of students' task processing (Aleven, McLaren, et al., 2016; Graesser et al., 2018; Koedinger & Corbett, 2005). Different reviews and meta-analyses reported that ITS outperform large-group instruction and can even achieve effects similar to human tutoring (Ma et al., 2014; Steenbergen-Hu & Cooper, 2014; VanLehn, 2011).

Researchers have considered a range of reasons potentially explaining the low robustness of ATI findings, especially compared to the larger effects of ITS (see Tetzlaff et al., 2021). One factor that is increasingly acknowledged is the versatile nature of some learner variables that, accordingly, cannot be reliably assessed with single-time measures. Therefore, different researchers have suggested to consider different timeframes for updating learner models and adapting learning experiences (Plass & Pawar, 2020; Tetzlaff et al., 2021): A macro-level strategy, as researched in ATI, can be employed to make adjustments based on learner variables and measures that are rather stable and relate to the wider learning context (Plass & Pawar, 2020). For example, macro-level adaptivity can adjust to aggregated measures of learner variables – such as prior knowledge (e.g., based on data from prior learning activities) or learning performance as an aggregate of recent learning behavior and outcomes (e.g., Nickl et al., 2022) – or to rather stable learner characteristics, such as learners' personality or general preferences (Rivers, 2021). Thereby, an adaptive learning environment can, for example, draw on overall performance measures (Bodily & Verbert, 2017) or suggest further resources and learning directions (Santos et al., 2014; Sevarac et al., 2012). In contrast, by adopting a micro-level strategy, real-time measurements of versatile learner variables, such as current learning behavior, can be used to monitor, adjust, and support the ongoing learning task (Plass & Pawar, 2020). Micro-level adaptivity can be found in many ITS and is particularly relevant for real-time learning support that adapts to learners needs during task processing.

There is a range of options for implementing adaptive learning support. In this article, we suggest distinguishing between three broad categories of learning support, namely, monitoring tools, instructional support, and task adjustment. The implementation of these types of learning support can further vary, depending on the decision-making agent – namely, teacher, learner, or the technology – that controls the activation, selection, and degree of the support. The adaptive educational technology further varies regarding the degree of automation versus human control in making assessments and deciding on interventions. For example, Molenaar (2022) suggested six levels of automation in the context of personalized teaching and learning, ranging from no technology support (teacher only), over varying degrees of shared responsibilities between the teacher and the technology, to fully automated personalization through technology. Besides these implementations of adaptive learning support, more general interventions, for example targeting learners' accessibility across learners and regions might be in focus of LA as well (Yan et al., 2024). However, and in line with our focus on fostering learning processes, we will focus on differentiating monitoring tools, instructional support, and task adjustment.

**2.3.2. Current Trends - Matching Educational Technologies and Theoretical Approaches**

Educational technology can facilitate teachers' or learners' monitoring and assessment of selected learner variables. We characterize such tools that summarize and present generated analytics about learner variables to facilitate teachers' or learners' monitoring as monitoring tools. This requires adaptive interfaces that visualize the LA generated by educational technologies. A popular type of monitoring tools are dashboards, which are customizable control panels that visualize LA generated by educational technologies in real time (e.g., traces of learning activities; Jivet et al., 2017; Sahin & Ifenthaler, 2021; Verbert et al., 2013). There are various educational technologies that can generate the data, such as tablets, online learning environments, virtual and mixed reality environments, or educational robots.

**Teacher dashboards** visualize analytics on relevant learner variables to facilitate teachers' monitoring and assessment, and ultimately, their decision-making about potential educational interventions. Knoop-van Campen et al. (2023) empirically investigated how a lesson overview dashboard showing real-time data of learners' performance influences teaching behaviors: They found that teachers using a dashboard rather gave feedback on learners' task processing instead of only giving feedback on the correctness of a task solution; this effect was especially pronounced for low-ability students. Moreover, Xhakaj et al. (2017) found that a teacher dashboard influences teachers' knowledge about their students, their lesson-planning, and the content that teachers' cover in class in a positive way. Many dashboards visualize rather summative learner metrics (e.g., learning progress, time on task, or error rates). But there are also attempts to consider theories of teaching and learning in dashboard designs: For example, the design of the teacher dashboard MetaDash is based on theoretical models and empirical evidence of self-regulated learning, aiming to support teachers' monitoring of students' self-regulated learning by visualizing relevant behavioral and process data of students (Wiedbusch et al., 2021). Besides teacher dashboards, some researchers also explored the use of smart glasses as **mixed-reality teacher monitoring tools** that show LA embedded within teachers' view of the classroom (e.g., alerting teachers to students that need help; Holstein et al., 2018). Although the majority of such technologies focuses on facilitating monitoring and assessment, educational technologies can also suggest potential interventions. For example, in the context of classroom orchestration, teachers might receive advice on how to pair students for collaborative learning tasks (Lawrence et al., 2023). Such types of support maintain the teachers' autonomy in deciding which intervention is best suited or whether intervening is necessary at all (see Molenaar, 2022). Learners can as well benefit from monitoring tools to enhance their awareness of their own learning prerequisites and processes (e.g., learning time and progress).

**Student facing learning dashboards** offer learners a visual overview of variables, such as their learning activities and outcomes (Jivet et al., 2017). For example, Kim et al. (2016) developed a dashboard for an online statistics course at a Korean university and tested its effect in comparison to a control group. They found that students who had access to a learning dashboard achieved higher final scores compared to those who did not. Sedrakyan et al. (2017) evaluated student-facing learning dashboards of affective states and found that students' awareness about their emotions during learning activities based on the visualization interpretation varied depending on previous knowledge on visualization techniques; generally, simpler visualizations resulted in better outcomes than more complex techniques. A systematic review of student-facing dashboards by Bodily & Verbert (2017) found that the majority of learning dashboards focuses on data visualizations (see also Verbert et al., 2013), although there are also few recommender systems (e.g., recommending further learning resources to students; Santos et al., 2014). The authors also reviewed the effects of learning dashboards on cognitive and behavioral learning outcomes and found mixed results, which they attributed to insufficient usability testing (Bodily & Verbert, 2017). Jivet et al. (2017) argue that especially learners with low self-regulation skills might experience challenges in using learning dashboards effectively and that dashboards often lack alignment with educational and psychological theory. However, more systematic research on the effects of learning dashboards on learners' cognitive outcomes as well as their non-cognitive outcomes is needed, as studies often only refer to affective outcomes in their evaluations (Valle et al., 2021). Besides learning dashboards, data can also be recorded, analyzed, and visualized using **tangible awareness tools**. For example, the CUBE, a cube-shaped tangible device with a screen at each side, can analyze and visualize learners' speaking time in small group discussions (Papadopoulos, 2019). It was argued that student facing monitoring tools represent powerful metacognitive resources for learners, triggering them to reason about their learning activities and outcomes (Charleer et al., 2016; Jivet et al., 2017). However, a systematic review of LA dashboards by Matcha et al. (2020) found that LA dashboards are rarely grounded in educational and psychological theory and often have significant limitations in how their evaluation was conducted and reported (e.g., lack of alignment between intended outcomes and measured variables; Valle et al., 2021). Thus, there is a need for further research extending the current state of theory-informed design of monitoring tools for teachers and learners (Baek & Doleck, 2023; Crompton et al., 2020; Jivet et al., 2017; Sahin & Ifenthaler, 2021; Wong et al., 2019): Promising candidates for theoretically grounded directions are self-regulated learning (Butler & Winne, 1995; Zimmerman, 2002), self-assessment (Andrade, 2019), internal feedback (Nicol, 2021), productive failure (Kapur & Rummel, 2012), and help-seeking (Aleven et al., 2003).

Besides students and teachers, technology itself can act as the agent in personalizing learning (i.e., adaptivity) – which can, however, again be regulated by human agents (i.e., adaptability; G. Fischer, 2001; Kucirkova et al., 2021; Plass & Pawar, 2020). Technology-based adaptivity is oftentimes used for automatic adaptive **instructional support**, that is, support measures that provide additional instruction besides the instruction that guides the core learning task.

**Feedback** is a typical instructional support measure that informs learners about their current state of performance (feed back) in relation to the learning goals (feed up) and highlights ways for the learner to move forward (feed forward), such as further resources or specific options for improving the previously shown performance (Hattie & Timperley, 2007; Wisniewski et al., 2020). Instructional feedback is different from the internal feedback generated by learners (which can be supported by monitoring and awareness tools) in that it provides instructional information to facilitate learners' task processing and performance. In a review of meta-analyses, Hattie & Timperley (2007), found that instructional feedback has medium to large effects on students' learning outcomes; however, the effect sizes varied considerably, which is why they discussed the quality of feedback as an essential factor and emphasized the role of personalized feedback that provides adaptive information tailored to a learner's needs. However, in online learning environments, there are various ways of how feedback is implemented (Bimba et al., 2017; Hattie & Timperley, 2007; Narciss et al., 2014; Wisniewski et al., 2020): non-adaptive feedback, such as presenting the correct response upon learners' submission of answers, is particularly easy to implement but does not provide any personalized information (e.g., Attali, 2015); simple adaptive feedback, such as providing an adaptive response whether the response was correct or incorrect (e.g., Stark et al., 2011), provides a summative assessment and, thus, provides little more information compared to monitoring tools; elaborated adaptive feedback provides not only information about the correctness of the solution but also involves a formative assessment of the learners' task processing. However, elaborated adaptive feedback often requires a more detailed analysis of learners' task processing and outcomes to identify, at what point or in which regard learners' task processing was flawed; this requires insight into learners' task processing and performance, for example, by analyzing learners' written responses. Zhu et al. (2017, 2020) used natural language processing for a formative feedback system integrated into an online science curriculum module on climate change and found that the feedback guided students' revisions of their solutions and improved their written scientific argumentation. Similarly, Sailer et al. (2022) used natural language processing to analyze learners' written responses in a simulation-based learning environment facilitating pre-service teachers' diagnostic reasoning and found a positive effect of the adaptive compared to static feedback on learners' quality in justifying their judgments.

Besides feedback, another prominent type of instructional support is **scaffolding**, that is, additional instruction that is given besides the instruction for the core learning task and provides guiding structures for task processing without which learners would not be able to achieve the same performance (Belland, 2014; Tabak & Kyza, 2018; Wood et al., 1976). Compared to feedback, which is more focused on evaluating learners' previous performance, scaffolding is more focused on providing instruction in how to optimize moving forward already during the task processing. There are several different types of scaffolding (e.g., Chernikova et al., 2020): For example, worked examples and modeling examples (Renkl, 2014; Van Gog & Rummel, 2010) exemplify how to solve a task; prompts and hints (e.g., cognitive and metacognitive prompts; Berthold et al., 2007; Quintana et al., 2004) give a more direct instruction regarding what aspects or steps of the task the learners might pay specific attention to; scripts and roles (F. Fischer et al., 2013; e.g., collaboration scripts; Vogel et al., 2017) guide the responsibilities and steps of task processing in further detail; moreover, reflection phases (Mamede & Schmidt,

2017) facilitate learners' assessment about their current performance (i.e., internal feedback) as well as their goal-setting, planning, and monitoring (i.e., self-regulation). Brush & Saye (2002) suggested to distinguish hard and soft scaffolding. Hard scaffolds are static supports that can be anticipated and planned in advance based upon typical learner difficulties with a task; as indicated by a meta-analysis, such non-adaptive scaffolding achieves medium-sized positive effects on cognitive outcomes in STEM education (Belland et al., 2017). By comparison, soft scaffolds provide dynamic, situation-specific help with the learning task and process (i.e., adaptive scaffolding). The relevance of adaptive scaffolding is highlighted by results from a meta-analysis indicating that different types of scaffolds are more or less effective for different types of learners, namely, worked examples were not effective for learners with high knowledge who benefitted most from reflection phases (Chernikova et al., 2020). The phenomenon that scaffolds can start to hinder learners' task processing and outcomes as their skill and autonomy increases was described as expertise reversal effect (Kalyuga et al., 2003). This effect is one of the reasons why several researchers have highlighted the role of fading as an integral part of scaffolding: Once learners become capable of doing a task on their own, the additional support structures should be gradually reduced and finally removed to further facilitate learners' autonomous task processing (Belland, 2014; Pea, 2004). Thus, similar to feedback, adaptivity also plays an important role for high-quality scaffolding, which is increasingly addressed in research. For example, Lim et al. (2023; see also van der Graaf et al., 2023 Van Der Graaf et al., 2023) investigated the effects of adaptive scaffolds in an online learning environment with reading and writing tasks and found a positive effect on SRL activities in the learning process. Radkowitsch et al. (2021) investigated adaptive collaboration scripts in an agent-based simulation in medical education and found positive effects on learners' collaboration activities and perceived competence.

Besides adapting instructional support to facilitate learners' performance of a learning task, educational technology can also automatically adjust the learning task itself. Automatic **task adjustment** consists in adaptations of instructions guiding the core learning task as well as adaptations of the task selection and learning paths, learning materials, or forms of interaction with the learning materials and the learning environment.

Examples can be found especially in ITS that personalize **task selection and learning paths** for individual students (besides providing adaptive instructional support; Graesser et al., 2018; Koedinger & Corbett, 2005). Most ITS focus on cognitive tutoring in learning domains with computationally well-defined problems (e.g., mathematics, physics, information technology; Anderson et al., 1995; Graesser et al., 2018). ITS are informed by theory and evidence from domain-specific education and from cognitive sciences (e.g., ACT-R, Anderson et al., 1997; knowledge space theory, Falmagne et al., 1990). The most common approach to designing ITS is a cognitive task analysis combined with cognitive modeling of learners: Cognitive task analysis aims to understand the knowledge, skills, and strategies required for performing well in a task domain (Anderson & Schunn, 2000); cognitive modeling of learners aims to generate a detailed description of the knowledge involved in learners' performance in a given task domain (e.g., strategies, problem-solving principles, knowledge of how to apply problem-solving principles to a specific problem; Aleven et al., 2010). More recent approaches have replaced cognitive modeling with example-tracing, that is, the author of the tutoring system demonstrates examples of problem-solving steps to generate an initial behavior graph that subsequently can be modified and annotated (Aleven, McLaren, et al., 2016). The task analysis and cognitive or example-based analysis are used to create a system in which learners' performance is used to constantly update the underlying adaptive learner model that the system creates and stores for the individual learner (Corbett, 2000), for example, to automatically select subsequent learning tasks. One of the most widely used tutoring systems is ALEKS (i.e., Assessment and Learning in Knowledge Spaces), which assesses students' current state of knowledge to personalize learning paths for the subjects of mathematics, chemistry, and statistics in secondary and post-secondary education (Falmagne et al., 2013). A meta-analysis found that the effectiveness of ALEKS is comparable with traditional classroom instruction, suggesting that the system can help teaching and learning, for example, by assisting classroom instruction or supporting students' homework (Fang et al., 2019). Besides focusing on cognitive support, some ITS have explored targeting other learner variables. For example, the approach of context personalization aims to spark situational interest by incorporating students' out-of-school interests into the learning tasks (e.g., mathematics tasks); a study by Bernacki & Walkington (2018) found that, compared to a non-personalized learning group, high school students in the personalized learning group reported higher situational interest as well as individual interest in mathematics and achieved higher exam performance. Other ITS have targeted learners' self-regulation, however, using adaptive instructional support (i.e., feedback and scaffolding) instead of task adjustment (e.g., Aleven et al., 2010; Azevedo et al., 2022).

In addition to task selection for personalizing students' learning paths, ITS have explored adjusting the forms of **interactions with the learning environment** by making use of conversational agents and chatbots. A prominent example is AutoTutor, which aims to simulate a human tutor (including an animated conversational agent) by engaging in a written conversation with the learner; AutoTutor was found to be effective in helping students learn science, technology, and other technical subject matters by interactively discussing questions and problems that require reasoning and explanations (Graesser, 2016; Graesser et al., 2005, 2012). Besides adapting to learners' cognitive activities and task-related content of written responses, some versions of the system explored adapting to affective reactions in learners' written responses, such as confusion, frustration, and boredom (D'mello & Graesser, 2012). Pedagogical conversational agents are considered a rising approach to learning and teaching (Alemdag, 2023; Schlimbach et al., 2022; Wollny et al., 2021). Especially with the rise of LLMs as accessible, for example, via ChatGPT (OpenAI, 2023), this development is increasingly gathering pace and might result in massive advancements in adaptive educational technologies (Kasneci et al., 2023). For example, the LLM GPT-3 was used as a pedagogical conversational agent to guide children's question-asking and spark their curiosity through curiosity-prompting cues for asking more and more profound questions (Abdelghani et al., 2022). However, despite the great opportunities offered by generative AI and LLMs, challenges and potential risks – such as ensuring ethics and data protection which is complicated through the opaque nature of deep learning algorithms – need to be handled with caution (Bauer et al., 2023; Conati et al., 2018; Drachsler & Greller, 2016; Kasneci et al., 2023).

Besides tutoring systems, adaptive task adjustment was also suggested as useful for other learning environments, such as digital simulations. Simulations are simplified but valid representations of natural, social or artificial systems, which include features that learners can manipulate (e.g., to approximate practice; Heitzmann et al., 2019; Sauvé et al., 2007). Designing simulations allows balancing authenticity and difficulty of the learning task (Codreanu et al., 2020), for example, by simplifying the simulated cases or situations, which is an example for adjusting **learning materials and interactions with the learning materials**. To systematize this form of task adjustments in simulations, a set of representational features has been suggested that can guide designing and adjusting (as well as sequencing, i.e., personalizing task selection and learning paths) simulated cases or situations – which has been characterized as representational scaffolding in simulations (F. Fischer et al., 2022). The representational features – namely, informational complexity, typicality, agency, and situation dynamics – and related scaffolds are grounded in educational and psychological theory (e.g., case-based reasoning, dual processing, complex problem-solving, cognitive load theory; Kolodner, 1992; Norman et al., 2007; Papa, 2016; Stadler et al., 2019; Sweller, 2010). However, the concept of representational scaffolding in simulations is a novel approach that is yet to be investigated by research. Nevertheless, as for ITS, using data-based analytics for task adjustments bears great potential for leveraging the effectiveness of the simulation-based approach for teaching and learning.

An overview of different adaptivity and personalization approaches sorted by their degree of task modification is summarized in Figure 2.

## 3. Discussion

### 3.1 Summary

The *Closed-Loop Learning Analytics Framework*, encompassing data collection, processing and analysis, and adaptivity and personalization, provides a comprehensive framework for understanding and enhancing the learning process. This framework not only aids in identifying research gaps and trends but also underscores the intricate interplay between technological innovation and educational as well as psychological theory in advancing the field of LA. Specifically, our review along the steps of our closed-loop framework shows significant advancements that have been made, particularly in data acquisition and processing. However, it also shows research gaps and lack of theoretical embedding in personalization of learning and a general lack of focus on learning processes and optimization of learning from an instructional design perspective, although suitable educational and psychological theories exist (see Section 2.3).

For the single steps of our closed-loop framework, the results of our review can be summarized as follows: The evolution of technology has significantly expanded the potential for data acquisition in LA. Contemporary methods go beyond traditional tools like questionnaires, embracing digital trace data, sensor-based collection, and biometrics. This technological advancement offers potential, especially when considering the combination of various data sources. Multimodal data acquisition has emerged as a particularly promising approach (Blikstein, 2013), aligning well with the complexity of the learning process. By integrating data from diverse sources, researchers can capture a more nuanced and holistic understanding of learners' interactions, behaviors, and cognitive states.

The field of LA has witnessed the development of a wide range of data processing and analytical methods, each with its unique advantages and challenges. Techniques such as supervised and unsupervised algorithms, relationship mining, and visualization have been employed to derive meaningful insights from complex datasets. The use of machine learning and AI technologies stands out as a significant trend (Kasneci et al., 2023), offering substantial potential for both current and future research. However, this advancement requires a critical balance between accuracy and interpretability (Giannakos & Cukurova, 2023; Hilbert et al., 2021). The diversity in methods reflects the varied nature of learning data and the need for tailored approaches to understand and predict learning behaviors and outcomes.

Adaptivity and personalization in LA have been identified as areas with significant research gaps, yet they hold many possibilities for enhancing learning experiences. The field has begun to explore various ways to adapt learning content, support, and environments based on individual learner needs, drawing on existing theories in psychology and education. However, a key question remains: What specific aspects of learning are we adapting to, and what exactly are we adapting, both on a micro and on a macro level – learning tools and environments, instructional methods, or learning tasks? (Bernacki et al., 2021). The concept of personalized learning involves tailoring instructional practices to the unique characteristics of each learner. This process entails both assessing relevant learner attributes (e.g., indicators for relevant cognitive, metacognitive, motivational-affective learner variables) and deciding on appropriate interventions (e.g., monitoring tools, instructional support, task adjustments). While educational and psychological theories offer some guidance, there is still a dire need for systematic approaches and robust frameworks to effectively implement personalized learning. The potential for real-time, dynamic adaptation presents both challenges and opportunities in creating more responsive and effective learning environments (Kasneci et al., 2023).

### 3.2 The Missing Link: Theory as a Core Connection

Learning has always been an area that has not only been driven by applied research in the classroom, but also by solid theories of learning and instruction closely related to educational and psychological knowledge on the underlying processes of learning, its antecedents, and its consequences. Educational and psychological theories act as a crucial connecting link across the various steps in LA – from data acquisition and processing to adaptivity and personalization. Educational and psychological theories not only provide coherence and direction to the research process (Crompton et al., 2020), but also balance the necessity of interpretability with the precision but also the dangers of purely data-driven insights. While we acknowledge that theories need to develop as new use cases and methods of data curation emerge, the theoretical embedding of LA has been underdeveloped leaving the field to some extent in a state of atheoretical fragmentation. A clear and dedicated focus on educational and psychological theory in applications of LA (see Figure 1) offers the chance to expand existing and develop new theories based on the advances that LA offer to the field of education and psychology and vice versa – theory expansions and developments that the fields are also in need of (Eronen & Bringmann, 2021; Renkl, 2023).

As described in the different steps in our closed-loop framework (see Section 2.1 - 2.3) and highlighted in reviews (e.g., Wong et al., 2019), there are certain psychological and educational theories that are being applied in LA research. Of these studies, many stem from the self-regulated learning context (Wong et al., 2019). In line with this focus, concepts of cognition, metacognition, and motivation/affect are in focus. From our perspective, cognitive theories, such as cognitive load theory (Sweller, 2024), offer valuable starting points for LA (O. Chen et al., 2023). In addition, theories with a focus on learning activities and learner products that are approximations for cognitive processes (e.g., ICAP model; Chi & Wylie, 2014) can inform decision making in different steps outlined in the closed loop. What is more, holistic approaches can benefit from combining different foci (e.g., cognition and affect) on aspects of learning as shown in the study by Taub et al. (2021) that is based on information processing theory (Winnie, 2017) and the model of affective dynamics (D'Mello & Grasser, 2012). Especially for personalization and adaptivity, instructional theories can obviously have a merit when designing instructional support (e.g., Belland, 2014; Hattie & Timperley, 2007).Theories serve as a navigational tool, guiding the orchestration of different stages in the LA process, which is vital for ensuring that each step aligns with the overarching educational goals and theoretical constructs.

On a more specific level, the current trend in LA research often oscillates between theory-driven and data-driven methodologies and sometimes does not clearly distinguish between micro- and macroadaptivity. While data-driven approaches offer granular insights and novel patterns, they sometimes lack the grounding in educational and psychological theory, leading to limited generalizability and practical implications beyond the actual context they are applied to (e.g., Gašević et al., 2015; Nistor et al., 2015; Viberg et al., 2018). In addition, a strong focus on educational and psychological theory might also enhance application of open science practices, such as preregistrations, in learning analytics research.

Placing theory at the center of the LA process – as done in our closed-loop – allows for 'loops within the loop'. This central positioning enables theory to bridge the gap between data collection, analysis, and application. It helps in conducting micro-level explorations within and between single steps while maintaining a cohesive macro-level perspective. The “loops within the loop” approach, wherein both application and theory development are considered, can be instrumental. By iteratively aligning practice with theory and vice versa, LA can evolve more consistently. This approach allows for continuous, fine-grained refinement of both theoretical frameworks and practical applications, leading to more robust and effective learning interventions.

Theoretical embedding, if adequately considered in advance, has strong implications for various steps in the loop as outlined in Figure 1. For instance, in data collection, the choice of what to measure (and what not to measure) should be guided by theoretical considerations. This is also in line with approaches emphasizing the measurement of relevant variables only, instead of all-in approaches for data collection (Drachsler, 2018). In fact, implicit assumptions and theories will guide data collection in any case, but making the underlying theoretical ideas and assumptions more explicit will be fundamental in establishing a strong link between educational and psychological theory and data-driven methodology. Further, theory-orientation in the operationalization of constructs can help to ensure generalizability that depends less on specific measures and manifest indicators. Similarly, in data processing and analysis, the selection of models and analysis techniques should align with theoretical requirements and the context of the study. In adaptivity and personalization interventions should be in line with established theories, considering the nuances of educational and psychological theory (Crompton et al., 2020) for the design of monitoring tools, instructional support, and task adjustments.

In conclusion, theory in LA acts as a critical link, ensuring that each step of the process is thoughtfully considered and aligned with educational objectives. By firmly rooting LA in educational and psychological theory, researchers can address existing challenges more effectively and pave the way for future innovations.

**3.3 Generative AI as a Chance and Challenge in Learning Analytics**

The advent of Generative AI, particularly LLMs, presents both significant opportunities and challenges in the realm of LA. These models have the potential to revolutionize how we approach learner models, formative assessment, and the development of advanced learning environments (see Kasneci et al., 2023). However, the effective application of these technologies hinges on the integration of solid theoretical foundations to guide their deployment and ensure meaningful learning experiences.

Generative AI has the potential to enable sophisticated learner models that can analyze and adapt to individual learning processes. They offer a promising avenue for conducting formative assessments in educational settings, processing and interpreting large volumes of textual data to provide immediate and contextually relevant feedback. This capability makes

LLMs a valuable tool for educators, facilitating continuous, adaptive feedback without the constraints of traditional methods (see Kasneci et al., 2023). Furthermore, the integration of LLMs into LA could lead to more interactive and adaptive learning experiences, ranging from personalized content delivery to intelligent tutoring systems (for a more comprehensive discussion on opportunities of generative AI in LA see Yan et al., 2024). However, the quality and effectiveness of these opportunities depend significantly on their alignment with educational objectives and theory.

Despite their potential, LLMs face challenges, particularly regarding transparency in their decision-making processes. This “black box” nature can impede understanding and trust in AI-driven learning tools. Moreover, the widespread availability of LLMs raises questions about the quality of learning opportunities they provide. Without a clear theoretical rationale, there's a risk of developing data-driven applications that do not effectively address learners' needs or fail to optimize learning processes. The theoretical grounding is essential for ensuring that adaptivity aligns with educational theories and pedagogical principles.

An emerging direction in the field is the development of alternative AI systems that require less computational power while performing equally well. For example, Rombach et al. (2022) present a viable solution for overcoming some of the limitations of current LLMs, particularly in terms of resource accessibility and efficiency. Such advancements could lead to more sustainable and accessible AI applications in LA, broadening the potential for their use in diverse educational settings.

In conclusion, while generative AI offers remarkable opportunities for advancing LA, its successful implementation requires careful consideration of both technological capabilities and theoretical underpinnings. The exploration of alternative, less resource-intensive AI systems opens new possibilities for creating effective, efficient, and accessible LA tools, further enhancing the field's potential to transform educational practices.

### 3.4 Conclusion

The exploration of LA from the perspective of the Closed-Loop Learning Analytics Framework outlined in this article underscores the field's complex and multifaceted nature. In reviewing approaches and trends in data collection, processing, analysis, and adaptivity, we emphasize the essential role of tenable educational and psychological theories as the backbone of LA research. However, to fully realize the potential of LA, embracing interdisciplinarity emerges as a crucial condition.

Interdisciplinarity in LA transcends the integration of data science, education, and psychology. It necessitates incorporating perspectives from diverse fields such as law, technology, and ethics to address the multifarious challenges and opportunities presented by advanced data technologies. Particularly with the advent of generative AI, such as LLMs, the interplay between technology and pedagogy becomes even more intricate. This complexity demands a broader, more inclusive approach to decision-making processes. These approaches should balance the efficiency and insights offered by AI systems with the nuanced understanding and ethical oversight provided by human expertise (Huber et al., 2024; Ninaus & Sailer, 2022). Such hybrid models could pave the way for more responsible, context-aware, and learner-centered applications of LA.

In conclusion, the future of LA lies in its ability to harmoniously blend data-driven insights with theoretical grounding and interdisciplinary perspectives. As we continue to navigate the evolving landscape of digital learning, a commitment to ethical principles, privacy considerations, and interdisciplinary collaboration will be instrumental in shaping learning environments that are not only effective and adaptive but also respectful of the diverse needs and rights of learners.

## 4. References


Abdelghani, R., Wang, Y.-H., Yuan, X., Wang, T., Lucas, P., Sauzéon, H., & Oudeyer, P.-Y. (2022). *GPT-3-driven pedagogical agents for training children's curious question-asking skills*. https://doi.org/10.48550/ARXIV.2211.14228

Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, *37*, 13–49. https://doi.org/10.1016/j.tele.2019.01.007

Alemdag, E. (2023). The effect of chatbots on learning: A meta-analysis of empirical research. *Journal of Research on Technology in Education*, 1–23. https://doi.org/10.1080/15391523.2023.2255698

Alemdag, E., & Yildirim, Z. (2022). Effectiveness of online regulation scaffolds on peer feedback provision and uptake: A mixed methods study. *Computers & Education*, *188*, 104574. https://doi.org/10.1016/j.compedu.2022.104574

Aleven, V., McLaren, B. M., Sewall, J., Van Velsen, M., Popescu, O., Demi, S., Ringenberg, M., & Koedinger, K. R. (2016). Example-Tracing Tutors: Intelligent Tutor Development for Non-programmers. *International Journal of Artificial Intelligence in Education*, *26*(1), 224–269. https://doi.org/10.1007/s40593-015-0088-2

Aleven, V., McLaughlin, E. A., Glenn, R. A., & Koedinger, K. R. (2016). Instruction based on adaptive learning technologies. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 522–560). Routledge. https://doi.org/10.4324/9781315736419.ch24

Aleven, V., Roll, I., McLAREN, B. M., & Koedinger, K. R. (2010). Automated, Unobtrusive, Action-by-Action Assessment of Self-Regulation During Learning With an Intelligent Tutoring System. *Educational Psychologist*, *45*(4), 224–233. https://doi.org/10.1080/00461520.2010.517740

Aleven, V., Stahl, E., Schworm, S., Fischer, F., & Wallace, R. (2003). Help Seeking and Help Design in Interactive Learning Environments. *Review of Educational Research*, *73*(3), 277–320. https://doi.org/10.3102/00346543073003277

Allen, L. K., Creer, S. C., & Öncel, P. (2022). Natural Language Processing: Towards a Multi-Dimensional View of the Learning Process. In C. Lang, G. Siemens, & A. F. Wise (Eds.), *The Handbook of Learning Analytics* (2nd ed., pp. 46–53). SOLAR. https://doi.org/10.18608/hla22.005

Alonso-Fernández, C., Calvo-Morata, A., Freire, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2019). Applications of data science to game learning analytics data: A systematic literature review. *Computers & Education*, *141*(June), 103612. https://doi.org/10.1016/j.compedu.2019.103612

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, Ray. (1995). Cognitive Tutors: Lessons Learned. *Journal of the Learning Sciences*, *4*(2), 167–207. https://doi.org/10.1207/s15327809jls0402_2

Anderson, J. R., Matessa, M., & Lebiere, C. (1997). ACT-R: A Theory of Higher Level Cognition and Its Relation to Visual Attention. *Human–Computer Interaction*, *12*(4), 439–462. https://doi.org/10.1207/s15327051hci1204_5

Anderson, J. R., & Schunn, C. (2000). Implications of the ACT-R learning theory: No magic bullets. In R. Glaser (Ed.), *Advances in instructional Psychology* (pp. 1–33). Routledge. https://doi.org/10.4324/9780203774526

Andrade, H. L. (2019). A Critical Review of Research on Student Self-Assessment. *Frontiers in Education*, *4*, 87. https://doi.org/10.3389/feduc.2019.00087

Appel, T., Gerjets, P., Hoffman, S., Moeller, K., Ninaus, M., Scharinger, C., Sevcenko, N., Wortha, F., & Kasneci, E. (2021). Cross-task and Cross-participant Classification of Cognitive Load in an Emergency Simulation Game. *IEEE Transactions on Affective Computing*, 1–1. https://doi.org/10.1109/TAFFC.2021.3098237

Arias, V. B., Ponce, F. P., & Martínez-Molina, A. (2023). How a Few Inconsistent Respondents Can Confound the Structure of Personality Survey Data: An Example With the Core-Self Evaluations Scale. *European Journal of Psychological Assessment*, *39*(6), 400–407. https://doi.org/10.1027/1015-5759/a000719

Attali, Y. (2015). Effects of multiple-try feedback and question type during mathematics problem solving on performance in similar problems. *Computers & Education*, *86*, 260–267. https://doi.org/10.1016/j.compedu.2015.08.011

Ayres, P., Lee, J. Y., Paas, F., & Van Merriënboer, J. J. G. (2021). The Validity of Physiological Measures to Identify Differences in Intrinsic Cognitive Load. *Frontiers in Psychology*, *12*, 702538. https://doi.org/10.3389/fpsyg.2021.702538

Azevedo, R., Bouchet, F., Duffy, M., Harley, J., Taub, M., Trevors, G., Cloude, E., Dever, D., Wiedbusch, M. D., Wortha, F., & Cerezo, R. (2022). Lessons Learned and Future Directions of MetaTutor: Leveraging Multichannel Data to Scaffold Self-Regulated Learning With an Intelligent Tutoring System. *Frontiers in Psychology*, *13*, 813632. https://doi.org/10.3389/fpsyg.2022.813632

Azevedo, R., & Gašević, D. (2019). Analyzing Multimodal Multichannel Data about Self-Regulated Learning with Advanced Learning Technologies: Issues and Challenges. *Computers in Human Behavior*, *96*, 207–210. https://doi.org/10.1016/j.chb.2019.03.025

Azevedo, R., Johnson, A., Chauncey, A., & Burkett, C. (2010). Self-regulated Learning with MetaTutor: Advancing the Science of Learning with MetaCognitive Tools. In M. S. Khine & I. M. Saleh (Eds.), *New Science of Learning* (pp. 225–247). Springer New York. https://doi.org/10.1007/978-1-4419-5716-0_11

Baek, C., & Doleck, T. (2023). Educational Data Mining versus Learning Analytics: A Review of Publications From 2015 to 2019. *Interactive Learning Environments*, *31*(6), 3828–3850. https://doi.org/10.1080/10494820.2021.1943689

Baker, R. S. (2023). Learning Analytics: An Opportunity for Education. *XRDS: Crossroads, The ACM Magazine for Students*, *29*(3), 18–21. https://doi.org/10.1145/3589636

Baker, R. S., & Siemens, G. (2022). Learning Analytics and Educational Data Mining. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (3rd ed., pp. 259–278). Cambridge University Press. https://doi.org/10.1017/9781108888295.016

Baker, R. S., & Yacef, K. (2009). The State of Educational Data Mining in 2009: A Review and Future Visions. *Journal of Educational Data Mining*, *1*(1), 3–17. https://doi.org/10.5281/ZENODO.3554658

Banihashem, S. K., Dehghanzadeh, H., Clark, D., Noroozi, O., & Biemans, H. J. A. (2023). Learning analytics for online game-Based learning: A systematic literature review. *Behaviour & Information Technology*, 1–28. https://doi.org/10.1080/0144929X.2023.2255301

Bauer, E., Greisel, M., Kuznetsov, I., Berndt, M., Kollar, I., Dresel, M., Fischer, M. R., & Fischer, F. (2023). Using natural language processing to support peer-feedback in the age of artificial intelligence: A cross-disciplinary framework and a research agenda. *British Journal of Educational Technology*, *54*(5), 1222–1245. https://doi.org/10.1111/bjet.13336

Bauer, E., Sailer, M., Kiesewetter, J., Fischer, M. R., & Fischer, F. (2022). Diagnostic argumentation in teacher education: Making the case for justification, disconfirmation, and transparency. *Frontiers in Education*, *7*, 977631. https://doi.org/10.3389/feduc.2022.977631

Belland, B. R. (2014). Scaffolding: Definition, Current Debates, and Future Directions. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology* (pp. 505–518). Springer New York. https://doi.org/10.1007/978-1-4614-3185-5_39

Belland, B. R., Walker, A. E., Kim, N. J., & Lefler, M. (2017). Synthesizing Results From Empirical Research on Computer-Based Scaffolding in STEM Education: A Meta-Analysis. *Review of Educational Research*, *87*(2), 309–344. https://doi.org/10.3102/0034654316670999

Bent, B., & Dunn, J. P. (2021). Optimizing sampling rate of wrist-worn optical sensors for physiologic monitoring. *Journal of Clinical and Translational Science*, *5*(1), e34. https://doi.org/10.1017/cts.2020.526

Bernacki, M. L., Greene, M. J., & Lobczowski, N. G. (2021). A Systematic Review of Research on Personalized Learning: Personalized by Whom, to What, How, and for What Purpose(s)? *Educational Psychology Review*, *33*(4), 1675–1715. https://doi.org/10.1007/s10648-021-09615-8

Bernacki, M. L., & Walkington, C. (2018). The role of situational interest in personalized learning. *Journal of Educational Psychology*, *110*(6), 864–881. https://doi.org/10.1037/edu0000250

Berndt, M., Strijbos, J.-W., & Fischer, F. (2018). Effects of written peer-feedback content and sender's competence on perceptions, performance, and mindful cognitive processing. *European Journal of Psychology of Education*, *33*(1), 31–49. https://doi.org/10.1007/s10212-017-0343-z

Berndt, M., Strijbos, J.-W., & Fischer, F. (2022). Impact of sender and peer-feedback characteristics on performance, cognitive load, and mindful cognitive processing. *Studies in Educational Evaluation*, *75*, 101197. https://doi.org/10.1016/j.stueduc.2022.101197

Bernecker, K., & Ninaus, M. (2021). No Pain, no Gain? Investigating motivational mechanisms of game elements in cognitive tasks. *Computers in Human Behavior*, *114*, 106542. https://doi.org/10.1016/j.chb.2020.106542

Berntson, G. G., Thomas Bigger, J., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., Nagaraja, H. N., Porges, S. W., Saul, J. P., Stone, P. H., & Van Der Molen, M. W. (1997). Heart rate variability: Origins, methods, and interpretive caveats. *Psychophysiology*, *34*(6), 623–648. https://doi.org/10.1111/j.1469-8986.1997.tb02140.x

Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction*, *17*(5), 564–577. https://doi.org/10.1016/j.learninstruc.2007.09.007

Bimba, A. T., Idris, N., Al-Hunaiyyan, A., Mahmud, R. B., & Shuib, N. L. B. M. (2017). Adaptive feedback in computer-based learning environments: A review. *Adaptive Behavior*, *25*(5), 217–234. https://doi.org/10.1177/1059712317727590

Blikstein, P. (2013). Multimodal learning analytics. *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, 102–106. https://doi.org/10.1145/2460296.2460316

Bodily, R., & Verbert, K. (2017). Trends and issues in student-facing learning analytics reporting systems research. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 309–318. https://doi.org/10.1145/3027385.3027403

Bolzer, M., Strijbos, J. W., & Fischer, F. (2015). Inferring mindful cognitive-processing of peer-feedback via eye-tracking: Role of feedback-characteristics, fixation-durations and transitions. *Journal of Computer Assisted Learning*, *31*(5), 422–434. https://doi.org/10.1111/jcal.12091

Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2023). *A Meta Systematic Review of Artificial Intelligence in Higher Education: A call for increased ethics, collaboration, and rigour*. https://doi.org/10.13140/RG.2.2.31921.56162/1

Borsboom, D., Mellenbergh, G. J., & Van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*(2), 203–219. https://doi.org/10.1037/0033-295X.110.2.203

Bosch, N. (2015). Multimodal Affect Detection in the Wild: Accuracy, Availability, and Generalizability. *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, 645–649. https://doi.org/10.1145/2818346.2823316

Bowers, A. J. (2010). Analyzing the Longitudinal K-12 Grading Histories of Entire Cohorts of Students: Grades, Data Driven Decision Making, Dropping Out and Hierarchical Cluster Analysis. *Practical Assessment, Research, and Evaluation*, *15*(1).

Braithwaite, J., Watson, D., Robert, J., & Mickey, R. (2013). *A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments* (pp. 1–42)

[Technical Report: Selective Attention & Awareness Laboratory (SAAL) Behavioural Brain Sciences Centre]. University of Birmingham.

Brush, T., & Saye, J. (2002). A summary of research exploring hard and soft scaffolding for teachers and students using a multimedia supported learning environment. *The Journal of Interactive Online Learning*, *1*(2), 1–12.

Butler, D. L., & Winne, P. H. (1995). Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research*, *65*(3), 245. https://doi.org/10.2307/1170684

Chaplot, D. S., Rhim, E., & Kim, J. (2015). Predicting Student Attrition in MOOCs using Sentiment Analysis and Neural Networks. *AIED 2015 Workshop Proceedings*, *3*.

Charleer, S., Klerkx, J., Duval, E., De Laet, T., & Verbert, K. (2016). Creating Effective Learning Analytics Dashboards: Lessons Learnt. In K. Verbert, M. Sharples, & T. Klobučar (Eds.), *Adaptive and Adaptable Learning* (Vol. 9891, pp. 42–56). Springer International Publishing. https://doi.org/10.1007/978-3-319-45153-4_4

Chen, O., Paas, F., & Sweller, J. (2023). A Cognitive Load Theory Approach to Defining and Measuring Task Complexity Through Element Interactivity. *Educational Psychology Review*, *35*(2), 63. https://doi.org/10.1007/s10648-023-09782-w

Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020). Application and theory gaps during the rise of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, *1*, 100002. https://doi.org/10.1016/j.caeai.2020.100002

Chernikova, O., Heitzmann, N., Fink, M. C., Timothy, V., Seidel, T., Fischer, F., & DFG Research group COSIMA. (2020). Facilitating Diagnostic Competences in Higher Education—A Meta-Analysis in Medical and Teacher Education. *Educational Psychology Review*, *32*(1), 157–196. https://doi.org/10.1007/s10648-019-09492-2

Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Cloude, E. B., Dever, D. A., Hahs-Vaughn, D. L., Emerson, A. J., Azevedo, R., & Lester, J. (2022). Affective Dynamics and Cognition During Game-Based Learning. *IEEE Transactions on Affective Computing*, *13*(4), 1705–1717. https://doi.org/10.1109/TAFFC.2022.3210755

Cloude, E. B., Dever, D. A., Wiedbusch, M. D., & Azevedo, R. (2020). Quantifying Scientific Thinking Using Multichannel Data With Crystal Island: Implications for Individualized Game-Learning Analytics. *Frontiers in Education*, *5*, 572546. https://doi.org/10.3389/feduc.2020.572546

Clow, D. (2012). The learning analytics cycle: Closing the loop effectively. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 134–138. https://doi.org/10.1145/2330601.2330636

Codreanu, E., Sommerhoff, D., Huber, S., Ufer, S., & Seidel, T. (2020). Between authenticity and cognitive demand: Finding a balance in designing a video-based simulation in the context of mathematics teacher education. *Teaching and Teacher Education*, *95*, 103146. https://doi.org/10.1016/j.tate.2020.103146

Conati, C., Porayska-Pomsta, K., & Mavrikis, M. (2018). *AI in Education needs interpretable machine learning: Lessons from Open Learner Modelling*. https://doi.org/10.48550/ARXIV.1807.00154

Corbett, A. (2000). *Cognitive mastery learning in the ACT programming tutor* (Adaptive User Interfaces. AAAI Technical Report SS-00-01). https://aaai.org/Library/Symposia/Spring/ss00-01.php

Crescenzi-Lanna, L. (2020). Multimodal Learning Analytics research with young children: A systematic review. *British Journal of Educational Technology*, *51*(5), 1485–1504. https://doi.org/10.1111/bjet.12959

Crompton, H., Bernacki, M., & Greene, J. A. (2020). Psychological foundations of emerging technologies for teaching and learning in higher education. *Current Opinion in Psychology*, *36*, 101–105. https://doi.org/10.1016/j.copsyc.2020.04.011

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*(11), 671–684. https://doi.org/10.1037/h0043943

Dalipi, F., Zdravkova, K., & Ahlgren, F. (2021). Sentiment Analysis of Students' Feedback in MOOCs: A Systematic Literature Review. *Frontiers in Artificial Intelligence*, *4*, 728708. https://doi.org/10.3389/frai.2021.728708

Di Mitri, D., Schneider, J., Specht, M., & Drachsler, H. (2018). From signals to knowledge: A conceptual model for multimodal learning analytics. *Journal of Computer Assisted Learning*, *34*(4), 338–349. https://doi.org/10.1111/jcal.12288

Dindar, M., Alikhani, I., Malmberg, J., Järvelä, S., & Seppänen, T. (2019). Examining shared monitoring in collaborative learning: A case of a recurrence quantification analysis approach. *Computers in Human Behavior, 100*, 335–344. https://doi.org/10.1016/j.chb.2019.03.004

D'mello, S., & Graesser, A. C. (2012). AutoTutor and affective autotutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Transactions on Interactive Intelligent Systems*, *2*(4), 1–39. https://doi.org/10.1145/2395123.2395128

D'Mello, S., & Graesser, A. C. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, *22*(2), 145–157. https://doi.org/10.1016/j.learninstruc.2011.10.001

Dowell, N., & Kovanović, V. (2022). Modeling Educational Discourse with Natural Language Processing. In C. Lang, G. Siemens, & A. F. Wise (Eds.), *The Handbook of Learning Analytics* (2nd ed., pp. 105–119). SOLAR. https://doi.org/10.18608/hla22.011

Drachsler, H. (2018). Trusted Learning Analytics. *Synergie*, *6*, 40–43. https://doi.org/10.25657/02:19141

Drachsler, H., & Greller, W. (2016). Privacy and Analytics – it's a DELICATE Issue A Checklist for Trusted Learning Analytics. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16*, 89–98. https://doi.org/10.1145/2883851.2883893

Drugova, E., Zhuravleva, I., Zakharova, U., & Latipov, A. (2023). Learning analytics driven improvements in learning design in higher education: A systematic literature review. *Journal of Computer Assisted Learning*, jcal.12894. https://doi.org/10.1111/jcal.12894

Du, X., Yang, J., Shelton, B. E., Hung, J.-L., & Zhang, M. (2021). A systematic meta-Review and analysis of learning analytics research. *Behaviour & Information Technology*, *40*(1), 49–62. https://doi.org/10.1080/0144929X.2019.1669712

Duan, L., Liu, J., Yang, W., Huang, T., & Gao, W. (2020). Video Coding for Machines: A Paradigm of Collaborative Compression and Intelligent Analytics. *IEEE Transactions on Image Processing*, *29*, 8680–8695. https://doi.org/10.1109/TIP.2020.3016485

Ehlenz, M., Heinemann, B., & Schroeder, U. (2022). Information Sources and their Potential for Multimodal Learning Analytics in Laboratory-based Learning. In D. Uckelmann, G. Romagnoli, J. Baalsrud Hauge, & V. Kammerlohr (Eds.), *Online-Labs in Education* (pp. 139–152). Nomos Verlagsgesellschaft mbH & Co. KG. https://doi.org/10.5771/9783957104106-139

Emerson, A., Cloude, E. B., Azevedo, R., & Lester, J. (2020). Multimodal learning analytics for game-based learning. *British Journal of Educational Technology*, *51*(5), 1505–1526. https://doi.org/10.1111/bjet.12992

Eronen, M. I., & Bringmann, L. F. (2021). The Theory Crisis in Psychology: How to Move Forward. *Perspectives on Psychological Science*, *16*(4), 779–788. https://doi.org/10.1177/1745691620970586

Falmagne, J.-C., Albert, D., Doble, C., Eppstein, D., & Hu, X. (Eds.). (2013). *Knowledge Spaces: Applications in Education*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-35329-1

Falmagne, J.-C., Koppen, M., Villano, M., Doignon, J.-P., & Johannesen, L. (1990). Introduction to knowledge spaces: How to build, test, and search them. *Psychological Review*, *97*(2), 201–224. https://doi.org/10.1037/0033-295X.97.2.201

Fan, Y., Van Der Graaf, J., Lim, L., Raković, M., Singh, S., Kilgour, J., Moore, J., Molenaar, I., Bannert, M., & Gašević, D. (2022). Towards investigating the validity of measurement of self-regulated learning based on trace data. *Metacognition and Learning*, *17*(3), 949–987. https://doi.org/10.1007/s11409-022-09291-1

Fang, Y., Ren, Z., Hu, X., & Graesser, A. C. (2019). A meta-analysis of the effectiveness of ALEKS on learning. *Educational Psychology*, *39*(10), 1278–1292. https://doi.org/10.1080/01443410.2018.1495829

Ferrari, M., & Quaresima, V. (2012). A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *NeuroImage*, *63*(2), 921–935. https://doi.org/10.1016/j.neuroimage.2012.03.049

Fischer, F., Bauer, E., Seidel, T., Schmidmaier, R., Radkowitsch, A., Neuhaus, B. J., Hofer, S. I., Sommerhoff, D., Ufer, S., Kuhn, J., Küchemann, S., Sailer, M., Koenen, J., Gartmeier, M., Berberat, P., Frenzel, A., Heitzmann, N., Holzberger, D., Pfeffer, J., … Fischer, M. R. (2022). Representational scaffolding in digital simulations – learning professional practices in higher education. *Information and Learning Sciences*, *123*(11/12), 645–665. https://doi.org/10.1108/ILS-06-2022-0076

Fischer, F., Kollar, I., Stegmann, K., & Wecker, C. (2013). Toward a Script Theory of Guidance in Computer-Supported Collaborative Learning. *Educational Psychologist*, *48*(1), 56–66. https://doi.org/10.1080/00461520.2012.748005

Fischer, G. (2001). User Modeling in Human–Computer Interaction. *User Modeling and User-Adapted Interaction*, *11*(1/2), 65–86. https://doi.org/10.1023/A:1011145532042

Fortenbacher, A., Ninaus, M., Yun, H., Helbig, R., & Moeller, K. (2019). Sensor Based Adaptive Learning—Lessons Learned. In N. Pinkwart & J. Konert (Eds.), *Die 17. Fachtagung Bildungstechnologien, Lecture Notes in Informatics (LNI), Gesellschaft für Informatik: Vol. P-297* (pp. 193–198). https://doi.org/10.18420/delfi2019_355

Freeman, J., Avons, S. E. S., Pearson, D. E. D., & Ijsselsteijn, W. A. (1999). Effects of Sensory Information and Prior Experience on Direct Subjective Ratings of Presence. *Presence: Teleoperators and Virtual Environments*, *8*(1), 1–13. https://doi.org/10.1162/105474699566017

Frenzel, A. C., Dindar, M., Pekrun, R., Reck, C., & Marx, A. K. G. (2024). Joy is reciprocally transmitted between teachers and students: Evidence on facial mimicry in the classroom. *Learning and Instruction*, *91*, 101896. https://doi.org/10.1016/j.learninstruc.2024.101896

Gal, R., May, A. M., van Overmeeren, E. J., Simons, M., & Monninkhof, E. M. (2018). The Effect of Physical Activity Interventions Comprising Wearables and Smartphone Applications on Physical Activity: A Systematic Review and Meta-analysis. *Sports Medicine - Open*, *4*(1), 42. https://doi.org/10.1186/s40798-018-0157-9

Gao, N., Shao, W., Rahaman, M. S., & Salim, F. D. (2020). n-Gage: Predicting in-class Emotional, Behavioural and Cognitive Engagement in the Wild. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *4*(3), 1–26. https://doi.org/10.1145/3411813

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, *59*(1), 64–71. https://doi.org/10.1007/s11528-014-0822-x

Giannakos, M. N., & Cukurova, M. (2023). The role of learning theory in multimodal learning analytics. *British Journal of Educational Technology*, *54*(5), 1246–1267. https://doi.org/10.1111/bjet.13320

Giannakos, M. N., Sharma, K., Papavlasopoulou, S., Pappas, I. O., & Kostakos, V. (2020). Fitbit for learning: Towards capturing the learning experience using wearable sensing. *International Journal of Human-Computer Studies*, *136*, 102384. https://doi.org/10.1016/j.ijhcs.2019.102384

Giannakos, M. N., Sharma, K., Pappas, I. O., Kostakos, V., & Velloso, E. (2019). Multimodal data as a means to understand the learning experience. *International Journal of Information Management*, *48*, 108–119. https://doi.org/10.1016/j.ijinfomgt.2019.02.003

Graesser, A. C. (2016). Conversations with AutoTutor Help Students Learn. *International Journal of Artificial Intelligence in Education*, *26*(1), 124–132. https://doi.org/10.1007/s40593-015-0086-4

Graesser, A. C., Chipman, P., Haynes, B. C., & Olney, A. (2005). AutoTutor: An Intelligent Tutoring System With Mixed-Initiative Dialogue. *IEEE Transactions on Education*, *48*(4), 612–618. https://doi.org/10.1109/TE.2005.856149

Graesser, A. C., D'Mello, S., Hu, X., Cai, Z., Olney, A., & Morgan, B. (2012). AutoTutor. In *Applied natural language processing: Identification, investigation and resolution* (pp. 169–187). IGI Global. https://doi.org/10.4018/978-1-60960-741-8.ch010

Graesser, A. C., Hu, X., & Sottilare, R. (2018). Intelligent Tutoring Systems. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Reimann (Eds.), *International Handbook of the Learning Sciences* (1st ed., pp. 246–255). Routledge. https://doi.org/10.4324/9781315617572-24

Gray, G., & Bergner, Y. (2022). A Practitioner's Guide to Measurement in Learning Analytics: Decisions, Opportunities, and Challenges. In C. Lang, G. Siemens, & A. F. Wise (Eds.), *The Handbook of Learning Analytics* (2nd ed., pp. 20–28). SOLAR. https://doi.org/10.18608/hla22.002

Greipl, S., Bernecker, K., & Ninaus, M. (2021). Facial and Bodily Expressions of Emotional Engagement: How Dynamic Measures Reflect the Use of Game Elements and Subjective Experience of Emotions and Effort. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CHI PLAY), 1–25. https://doi.org/10.1145/3474667

Greipl, S., Klein, E., Lindstedt, A., Kiili, K., Moeller, K., Karnath, H.-O., Bahnmueller, J., Bloechle, J., & Ninaus, M. (2021). When the brain comes into play: Neurofunctional correlates of emotions and reward in game-based learning. *Computers in Human Behavior, 125*, 106946. https://doi.org/10.1016/j.chb.2021.106946

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Heitzmann, N., Seidel, T., Hetmanek, A., Wecker, C., Fischer, M. R., Ufer, S., Schmidmaier, R., Neuhaus, B., Siebeck, M., Stürmer, K., Obersteiner, A., Reiss, K., Girwidz, R., Fischer, F., & Opitz, A. (2019). Facilitating Diagnostic Competences in Simulations in Higher Education A Framework and a Research Agenda. *Frontline Learning Research*, 1–24. https://doi.org/10.14786/flr.v7i4.384

Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education*, *90*(December 2015), 36–53. https://doi.org/10.1016/j.compedu.2015.09.005

Hershkovitz, A., De Baker, R. S. J., Gobert, J., Wixon, M., & Pedro, M. S. (2013). Discovery With Models: A Case Study on Carelessness in Computer-Based Science Inquiry. *American Behavioral Scientist*, *57*(10), 1480–1499. https://doi.org/10.1177/0002764213479365

Hilbert, S., Coors, S., Kraus, E., Bischl, B., Lindl, A., Frei, M., Wild, J., Krauss, S., Goretzko, D., & Stachl, C. (2021). Machine learning for the educational sciences. *Review of Education*, *9*(3). https://doi.org/10.1002/rev3.3310

Holstein, K., McLaren, B. M., & Aleven, V. (2017). SPACLE: Investigating learning across virtual and physical spaces using spatial replays. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 358–367. https://doi.org/10.1145/3027385.3027450

Holstein, K., McLaren, B. M., & Aleven, V. (2018). Student Learning Benefits of a Mixed-Reality Teacher Awareness Tool in AI-Enhanced Classrooms. In C. Penstein Rosé, R. Martínez-Maldonado, H. U. Hoppe, R. Luckin, M. Mavrikis, K. Porayska-Pomsta, B. McLaren, & B. Du Boulay (Eds.), *Artificial Intelligence in Education* (Vol. 10947, pp. 154–168). Springer International Publishing. https://doi.org/10.1007/978-3-319-93843-1_12

Huber, S. E., Kiili, K., Nebel, S., Ryan, R. M., Sailer, M., & Ninaus, M. (2024). Leveraging the Potential of Large Language Models in Education Through Playful and Game-Based Learning. *Educational Psychology Review*, *36*(1), 25. https://doi.org/10.1007/s10648-024-09868-z

Huettel, S. A., Song, A. W., & McCarthy, G. (2009). *Functional Magnetic Resonance Imaging* (2nd ed.). Sinauer Associates, Inc.

Huisman, B., Saab, N., Van Driel, J., & Van Den Broek, P. (2020). A questionnaire to assess students' beliefs about peer-feedback. *Innovations in Education and Teaching International*, *57*(3), 328–338. https://doi.org/10.1080/14703297.2019.1630294

Hutt, S., Wong, A., Papoutsaki, A., Baker, R. S., Gold, J. I., & Mills, C. (2023). Webcam-based eye tracking to detect mind wandering and comprehension errors. *Behavior Research Methods*. https://doi.org/10.3758/s13428-022-02040-x

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning* (2nd ed.). Springer.

Janssen, T. W. P., Grammer, J. K., Bleichner, M. G., Bulgarelli, C., Davidesco, I., Dikker, S., Jasińska, K. K., Siugzdaite, R., Vassena, E., Vatakis, A., Zion-Golumbic, E., & Atteveldt, N. (2021). Opportunities and Limitations of Mobile Neuroimaging Technologies in Educational Neuroscience. *Mind, Brain, and Education*, *15*(4), 354–370. https://doi.org/10.1111/mbe.12302

Järvenoja, H., Järvelä, S., Törmänen, T., Näykki, P., Malmberg, J., Kurki, K., Mykkänen, A., & Isohätälä, J. (2018). Capturing motivation and emotion regulation during a learning process. *Frontline Learning Research*, 85–104. https://doi.org/10.14786/flr.v6i3.369

Jivet, I., Scheffel, M., Drachsler, H., & Specht, M. (2017). Awareness Is Not Enough: Pitfalls of Learning Analytics Dashboards in the Educational Practice. In É. Lavoué, H. Drachsler, K. Verbert, J. Broisin, & M. Pérez-Sanagustín (Eds.), *Data Driven Approaches in Digital Education* (Vol. 10474, pp. 82–96). Springer International Publishing. https://doi.org/10.1007/978-3-319-66610-5_7

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, *19*(4), 509–539. https://doi.org/10.1007/s10648-007-9054-3

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The Expertise Reversal Effect. *Educational Psychologist*, *38*(1), 23–31. https://doi.org/10.1207/S15326985EP3801_4

Kapur, M., & Rummel, N. (2012). Productive failure in learning from generation and invention activities. *Instructional Science*, *40*(4), 645–650. https://doi.org/10.1007/s11251-012-9235-4

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Khosravi, S., Khan, A. R., Zoha, A., & Ghannam, R. (2022). Self-Directed Learning using Eye-Tracking: A Comparison between Wearable Head-worn and Webcam-based Technologies. *2022 IEEE Global Engineering Education Conference (EDUCON)*, 640–643. https://doi.org/10.1109/EDUCON52537.2022.9766468

Kim, J., Jo, I.-H., & Park, Y. (2016). Effects of learning analytics dashboard: Analyzing the relations among dashboard utilization, satisfaction, and learning achievement. *Asia Pacific Education Review*, *17*(1), 13–24. https://doi.org/10.1007/s12564-015-9403-8

Kitto, K., Hicks, B., & Buckingham Shum, S. (2023). Using causal models to bridge the divide between big data and educational theory. *British Journal of Educational Technology*, *54*(5), 1095–1124. https://doi.org/10.1111/bjet.13321

Knoop-van Campen, C. A. N., Wise, A., & Molenaar, I. (2023). The equalizing effect of teacher dashboards on feedback in K-12 classrooms. *Interactive Learning Environments*, *31*(6), 3447–3463. https://doi.org/10.1080/10494820.2021.1931346

Kober, S. E., Ninaus, M., Friedrich, E. V. C., & Scherer, R. (2018). BCI and Games: Playful, Experience-Oriented Learning by Vivid Feedback? In C. S. Nam, A. Nijholt, & F. Lotte (Eds.), *Brain–Computer Interfaces Handbook: Technological and Theoretical Advances*. CRC Press - Taylor & Francis Group.

Koedinger, K., & Corbett, A. T. (2005). Cognitive Tutors: Technology bringing learning science to the classroom. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 61–78). Cambridge University Press.

Kolodner, J. L. (1992). An introduction to case-based reasoning. *Artificial Intelligence Review*, *6*(1), 3–34. https://doi.org/10.1007/BF00155578

Krenn, M., Pollice, R., Guo, S. Y., Aldeghi, M., Cervera-Lierta, A., Friederich, P., dos Passos Gomes, G., Häse, F., Jinich, A., Nigam, A., Yao, Z., & Aspuru-Guzik, A. (2022). On scientific understanding with artificial intelligence. *Nature Reviews Physics*, *4*(12), 761–769. https://doi.org/10.1038/s42254-022-00518-3

Kucirkova, N., Gerard, L., & Linn, M. C. (2021). Designing personalised instruction: A research and design framework. *British Journal of Educational Technology*, *52*(5), 1839–1861. https://doi.org/10.1111/bjet.13119

Lawrence, L., Echeverria, V., Yang, K., Aleven, V., & Rummel, N. (2023). How teachers conceptualise shared control with an AI co-orchestration tool: A multiyear teacher-centred design process. *British Journal of Educational Technology*, bjet.13372. https://doi.org/10.1111/bjet.13372

Leifheit, L., Tsarava, K., Ninaus, M., Ostermann, K., Golle, J., Trautwein, U., & Moeller, K. (2020). SCAPA : Development of a Questionnaire Assessing Self-Concept and Attitudes Toward Programming. *Proceedings of the 25th Annual Conference on Innovation and Technology in Computer Science Education - ITiCSE 2020*, 138–144. https://doi.org/10.1145/3341525.3387415

Leitner, P., Khalil, M., & Ebner, M. (2017). Learning Analytics in Higher Education—A Literature Review. In A. Peña-Ayala (Ed.), *Learning Analytics: Fundaments, Applications, and Trends* (Vol. 94, pp. 1–23). Springer International Publishing. https://doi.org/10.1007/978-3-319-52977-6_1

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. *Proceedings of the 34th International Conference on Neural Information Processing Systems (NIPS'20)*, 9459–9474.

Li, Q., Baker, R., & Warschauer, M. (2020). Using clickstream data to measure, understand, and support self-regulated learning in online courses. *The Internet and Higher Education*, *45*, 100727. https://doi.org/10.1016/j.iheduc.2020.100727

Lim, L., Bannert, M., Van Der Graaf, J., Singh, S., Fan, Y., Surendrannair, S., Rakovic, M., Molenaar, I., Moore, J., & Gašević, D. (2023). Effects of real-time analytics-based personalized scaffolds on students' self-regulated learning. *Computers in Human Behavior*, *139*, 107547. https://doi.org/10.1016/j.chb.2022.107547

Long, P., & Siemens, G. (2011). Penetrating the Fog: Analytics in Learning and Education. *EDUCAUSE Review*, *46*(5).

Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*, *106*(4), 901–918. https://doi.org/10.1037/a0037123

Mamede, S., & Schmidt, H. G. (2017). Reflection in Medical Diagnosis: A Literature Review. *Health Professions Education*, *3*(1), 15–25. https://doi.org/10.1016/j.hpe.2017.01.003

Martin, P. P., Kranz, D., Wulff, P., & Graulich, N. (2023). Exploring new depths: Applying machine learning for the analysis of student argumentation in chemistry. *Journal of Research in Science Teaching*, tea.21903. https://doi.org/10.1002/tea.21903

Matcha, W., Gašević, D., Jovanović, J., Uzir, N. A., Oliver, C. W., Murray, A., & Gasevic, D. (2020). Analytics of learning strategies: The association with the personality traits. *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, 151–160. https://doi.org/10.1145/3375462.3375534

Matcha, W., Uzir, N. A., Gasevic, D., & Pardo, A. (2020). A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective. *IEEE Transactions on Learning Technologies*, *13*(2), 226–245. https://doi.org/10.1109/TLT.2019.2916802

Mills, C., Fridman, I., Soussou, W., Waghray, D., Olney, A. M., & D'Mello, S. K. (2017). Put your thinking cap on: Detecting cognitive load using EEG during learning. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 80–89. https://doi.org/10.1145/3027385.3027431

Molenaar, I. (2022). Towards hybrid human-AI learning technologies. *European Journal of Education*, *57*(4), 632–645. https://doi.org/10.1111/ejed.12527

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

Motz, B., Bergner, Y., Brooks, C. A., Gladden, A., Gray, G., Lang, C., Li, W., Marmolejo-Ramos, F., & Quick, J. D. (2023). LAK of Direction: Misalignment Between the Goals of Learning Analytics and its Research Scholarship. *Journal of Learning Analytics*, 1–13. https://doi.org/10.18608/jla.2023.7913

Motz, B., Quick, J., Schroeder, N., Zook, J., & Gunkel, M. (2019). The validity and utility of activity logs as a measure of student engagement. *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 300–309. https://doi.org/10.1145/3303772.3303789

Namoun, A., & Alshanqiti, A. (2020). Predicting Student Performance Using Data Mining and Learning Analytics Techniques: A Systematic Literature Review. *Applied Sciences*, *11*(1), 237. https://doi.org/10.3390/app11010237

Narciss, S., Sosnovsky, S., Schnaubert, L., Andrès, E., Eichelmann, A., Goguadze, G., & Melis, E. (2014). Exploring feedback and student characteristics relevant for personalizing feedback strategies. *Computers & Education*, *71*, 56–76. https://doi.org/10.1016/j.compedu.2013.09.011

Nebel, S., & Ninaus, M. (2019). New Perspectives on Game-Based Assessment with Process Data and Physiological Signals. In D. Ifenthaler & Y. Kim (Eds.), *Game-Based Assessment Revisited. Advances in Game-Based Learning* (pp. 141–161). Springer. https://doi.org/10.1007/978-3-030-15569-8_8

Nelson, L. K. (2020). Computational Grounded Theory: A Methodological Framework. *Sociological Methods & Research*, *49*(1), 3–42. https://doi.org/10.1177/0049124117729703

Nickl, M., Huber, S. A., Sommerhoff, D., Codreanu, E., Ufer, S., & Seidel, T. (2022). Video-based simulations in teacher education: The role of learner characteristics as capacities for positive learning experiences and high performance. *International Journal of Educational Technology in Higher Education*, *19*(1), 45. https://doi.org/10.1186/s41239-022-00351-9

Nicol, D. (2021). The power of internal feedback: Exploiting natural comparison processes. *Assessment & Evaluation in Higher Education*, *46*(5), 756–778. https://doi.org/10.1080/02602938.2020.1823314

Ninaus, M., Greipl, S., Kiili, K., Lindstedt, A., Huber, S., Klein, E., Karnath, H.-O., & Moeller, K. (2019). Increased emotional engagement in game-based learning – A machine learning approach on facial emotion detection data. *Computers & Education*, *142*(July), 103641. https://doi.org/10.1016/j.compedu.2019.103641

Ninaus, M., Kober, S. E., Friedrich, E. V. C., Dunwell, I., Freitas, S. D., Arnab, S., Ott, M., Kravcik, M., Lim, T., Louchart, S., Bellotti, F., Hannemann, A., Thin, A. G., Berta, R., Wood, G., & Neuper, C. (2014). Neurophysiological methods for monitoring brain activity in serious games and virtual environments: A review. *International Journal of Technology Enhanced Learning*, *6*(1), 78. https://doi.org/10.1504/IJTEL.2014.060022

Ninaus, M., Kober, S. E., Friedrich, E. V. C. E. V. C. E. V. C., Neuper, C., & Wood, G. (2014). The Potential Use of Neurophysiological Signals for Learning Analytics. *2014 6th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)*, 1–5. https://doi.org/10.1109/VS-Games.2014.7012169

Ninaus, M., & Nebel, S. (2021). A Systematic Literature Review of Analytics for Adaptivity Within Educational Video Games. *Frontiers in Education*, *5*, 611072. https://doi.org/10.3389/feduc.2020.611072

Ninaus, M., & Sailer, M. (2022). Closing the loop – The human role in artificial intelligence for education. *Frontiers in Psychology*, *13*, 956798. https://doi.org/10.3389/fpsyg.2022.956798

Ninaus, M., Tsarava, K., & Moeller, K. (2019). A Pilot Study on the Feasibility of Dynamic Difficulty Adjustment in Game-Based Learning Using Heart-Rate. In A. Liapis, G. N. Yannakakis, M. Gentile, & M. Ninaus (Eds.), *Games and Learning Alliance. GALA 2019. Lecture Notes in Computer Science* (Vol. 2, pp. 117–128). Springer International Publishing. https://doi.org/10.1007/978-3-030-34350-7_12

Ninaus, M., Witte, M., Kober, S. E., Friedrich, E. V. C., Kurzmann, J., Hartsuiker, E., Neuper, C., & Wood, G. (2013). Neurofeedback and Serious Games. In T. M. Connolly, E. Boyle, T. Hainey, G. Baxter, & P. Moreno-ger (Eds.), *Psychology, Pedagogy, and Assessment in Serious Games* (Vol. i, pp. 82–110). IGI Global. https://doi.org/10.4018/978-1-4666-4773-2.ch005

Nistor, N., Derntl, M., & Klamma, R. (2015). Learning Analytics: Trends and Issues of the Empirical Research of the Years 2011–2014. In G. Conole, T. Klobučar, C. Rensing, J. Konert, & E. Lavoué (Eds.), *Design for Teaching and Learning in a Networked World* (Vol. 9307, pp. 453–459). Springer International Publishing. https://doi.org/10.1007/978-3-319-24258-3_39

Norman, G., Young, M., & Brooks, L. (2007). Non-analytical models of clinical reasoning: The role of experience. *Medical Education*, *0*(0), 071116225013001-??? https://doi.org/10.1111/j.1365-2923.2007.02914.x

Ocak, C., Kopcha, T. J., & Dey, R. (2023). An AI-enhanced pattern recognition approach to temporal and spatial analysis of children's embodied interactions. *Computers and Education: Artificial Intelligence*, *5*, 100146. https://doi.org/10.1016/j.caeai.2023.100146

Ochoa, X. (2022). Multimodal Learning Analytics: Rationale, Process, Examples, and Direction. In C. Lang, G. Siemens, & A. F. Wise (Eds.), *The Handbook of Learning Analytics* (2nd ed., pp. 54–65). SOLAR. https://doi.org/10.18608/hla22.006

Omarchevska, Y., Lachner, A., Richter, J., & Scheiter, K. (2022). It takes two to tango: How scientific reasoning and self-regulation processes impact argumentation quality. *Journal of the Learning Sciences*, *31*(2), 237–277. https://doi.org/10.1080/10508406.2021.1966633

Omer, U., Tehseen, R., Farooq, M. S., & Abid, A. (2023). Learning analytics in programming courses: Review and implications. *Education and Information Technologies*, *28*(9), 11221–11268. https://doi.org/10.1007/s10639-023-11611-0

Ouhaichi, H., Spikol, D., & Vogel, B. (2023). Research trends in multimodal learning analytics: A systematic mapping study. *Computers and Education: Artificial Intelligence*, *4*, 100136. https://doi.org/10.1016/j.caeai.2023.100136

Ouyang, F., Wu, M., Zhang, L., Xu, W., Zheng, L., & Cukurova, M. (2023). Making strides towards AI-supported regulation of learning in collaborative knowledge construction. *Computers in Human Behavior*, *142*, 107650. https://doi.org/10.1016/j.chb.2023.107650

Papa, F. J. (2016). A Dual Processing Theory Based Approach to Instruction and Assessment of Diagnostic Competencies. *Medical Science Educator*, *26*(4), 787–795. https://doi.org/10.1007/s40670-016-0326-8

Papadopoulos, P. (2019). The CUBE: A Tangible for Embodied Learning, Balanced Engagement, and Classroom Orchestration. In K. Lund, G. Niccolai, E. Lavoué, C. E. Hmelo-Silver, G. Gweon, & M. Baker (Eds.), *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings, 13th International Conference on Computer Supported Collaborative Learning (CSCL) 2019* (Vol. 2, pp. 1001–1004). https://linkinghub.elsevier.com/retrieve/pii/S2212868917300995

Papamitsiou, Z., Pappas, I. O., Sharma, K., & Giannakos, M. N. (2020). Utilizing Multimodal Data Through fsQCA to Explain Engagement in Adaptive Learning. *IEEE Transactions on Learning Technologies*, *13*(4), 689–703. https://doi.org/10.1109/TLT.2020.3020499

Parry, D. A., Davidson, B. I., Sewall, C. J. R., Fisher, J. T., Mieczkowski, H., & Quintana, D. S. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-021-01117-5

Patchan, M. M., & Schunn, C. D. (2015). Understanding the benefits of providing peer feedback: How students respond to peers' texts of varying quality. *Instructional Science*, *43*(5), 591–614. https://doi.org/10.1007/s11251-015-9353-x

Pea, R. D. (2004). The Social and Technological Dimensions of Scaffolding and Related Theoretical Concepts for Learning, Education, and Human Activity. *Journal of the Learning Sciences*, *13*(3), 423–451. https://doi.org/10.1207/s15327809jls1303_6

Peddycord-Liu, Z., Harred, R., Karamarkovich, S., Barnes, T., Lynch, C., & Rutherford, T. (2018). Learning Curve Analysis in a Large-Scale, Drill-and-Practice Serious Math Game: Where Is

Learning Support Needed? In C. Penstein Rosé, R. Martínez-Maldonado, H. U. Hoppe, R. Luckin, M. Mavrikis, K. Porayska-Pomsta, B. McLaren, & B. Du Boulay (Eds.), *Artificial Intelligence in Education* (Vol. 10947, pp. 436–449). Springer International Publishing. https://doi.org/10.1007/978-3-319-93843-1_32

Pekrun, R., Vogl, E., Muis, K. R., & Sinatra, G. M. (2017). Measuring emotions during epistemic activities: The Epistemically-Related Emotion Scales. *Cognition and Emotion*, *31*(6), 1268–1276. https://doi.org/10.1080/02699931.2016.1204989

Plass, J. L., & Pawar, S. (2020). Toward a taxonomy of adaptivity for learning. *Journal of Research on Technology in Education*, *52*(3), 275–300. https://doi.org/10.1080/15391523.2020.1719943

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., Kyza, E., Edelson, D., & Soloway, E. (2004). A Scaffolding Design Framework for Software to Support Science Inquiry. *Journal of the Learning Sciences*, *13*(3), 337–386. https://doi.org/10.1207/s15327809jls1303_4

Radkowitsch, A., Sailer, M., Schmidmaier, R., Fischer, M. R., & Fischer, F. (2021). Learning to diagnose collaboratively – Effects of adaptive collaboration scripts in agent-based medical simulations. *Learning and Instruction*, *75*, 101487. https://doi.org/10.1016/j.learninstruc.2021.101487

Radkowitsch, A., Sommerhoff, D., Nickl, M., Codreanu, E., Ufer, S., & Seidel, T. (2023). Exploring the diagnostic process of pre-service teachers using a simulation – A latent profile approach. *Teaching and Teacher Education*, *130*, 104172. https://doi.org/10.1016/j.tate.2023.104172

Renkl, A. (2014). Toward an Instructionally Oriented Theory of Example-Based Learning. *Cognitive Science*, *38*(1), 1–37. https://doi.org/10.1111/cogs.12086

Renkl, A. (2023). Lernprozesse und deren Förderung theoretisch verstehen: Von kleinen Fortschritten, von gegenwärtig priorisierten, aber nicht ausreichenden Abhilfen und von einem Wunschzettel. *Unterrichtswissenschaft*, *51*(1), 39–62. https://doi.org/10.1007/s42010-023-00167-0

Riemer, V., Frommel, J., Layher, G., Neumann, H., & Schrader, C. (2017). Identifying features of bodily expression as indicators of emotional experience during multimedia learning. *Frontiers in Psychology*, *8*(JUL). https://doi.org/10.3389/fpsyg.2017.01303

Rivers, D. J. (2021). The role of personality traits and online academic self-efficacy in acceptance, actual use and achievement in Moodle. *Education and Information Technologies*, *26*(4), 4353–4378. https://doi.org/10.1007/s10639-021-10478-3

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-Resolution Image Synthesis with Latent Diffusion Models. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 10674–10685. https://doi.org/10.1109/CVPR52688.2022.01042

Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *WIREs Data Mining and Knowledge Discovery*, *10*(3), e1355. https://doi.org/10.1002/widm.1355

Sahin, M., & Ifenthaler, D. (2021). Visualizations and Dashboards for Learning Analytics: A Systematic Literature Review. In M. Sahin & D. Ifenthaler (Eds.), *Visualizations and Dashboards for Learning Analytics* (pp. 3–22). Springer International Publishing. https://doi.org/10.1007/978-3-030-81222-5_1

Sailer, M., Bauer, E., Hofmann, R., Kiesewetter, J., Glas, J., Gurevych, I., & Fischer, F. (2022). Adaptive feedback from artificial neural networks facilitates pre-service teachers’ diagnostic reasoning in simulation-based learning. *Learning and Instruction*, 101620. https://doi.org/10.1016/j.learninstruc.2022.101620

Samuelsen, J., Chen, W., & Wasson, B. (2019). Integrating multiple data sources for learning analytics—Review of literature. *Research and Practice in Technology Enhanced Learning*, *14*(1), 11. https://doi.org/10.1186/s41039-019-0105-4

Santos, O. C., Boticario, J. G., & Pérez-Marín, D. (2014). Extending web-based educational systems with personalised support through User Centred Designed recommendations along the e-learning life cycle. *Science of Computer Programming*, *88*, 92–109. https://doi.org/10.1016/j.scico.2013.12.004

Sauvé, L., Renaud, L., Kaufman, D., & Marquis, J.-S. (2007). Distinguishing between games and simulations: A systematic review. *Journal of Educational Technology & Society*, *10*(3), 247–256.

Schlimbach, R., Rinn, H., Markgraf, D., & Robra-Bissantz, S. (2022). A Literature Review on Pedagogical Conversational Agent Adaptation. *PACIS 2022 Proceedings*. https://aisel.aisnet.org/pacis2022/20

Schneider, J., Börner, D., Van Rosmalen, P., & Specht, M. (2015). Augmenting the senses: A review on sensor-based learning support. *Sensors (Switzerland)*, *15*(2), 4097–4133. https://doi.org/10.3390/s150204097

Sedrakyan, G., Leony, D., Muñoz-Merino, P. J., Kloos, C. D., & Verbert, K. (2017). Evaluating Student-Facing Learning Dashboards of Affective States. In É. Lavoué, H. Drachsler, K. Verbert, J. Broisin, & M. Pérez-Sanagustín (Eds.), *Data Driven Approaches in Digital Education* (Vol. 10474, pp. 224–237). Springer International Publishing. https://doi.org/10.1007/978-3-319-66610-5_17

Selwyn, N. (2019). What's the Problem with Learning Analytics? *Journal of Learning Analytics*, *6*(3). https://doi.org/10.18608/jla.2019.63.3

Sevarac, Z., Devedzic, V., & Jovanovic, J. (2012). Adaptive neuro-fuzzy pedagogical recommender. *Expert Systems with Applications*, *39*(10), 9797–9806. https://doi.org/10.1016/j.eswa.2012.02.174

Shaffer, D. W. (2017). *Quantitative ethnography*. Cathcart Press.

Sharma, K., & Giannakos, M. (2020). Multimodal data capabilities for learning: What can multimodal data tell us about learning? *British Journal of Educational Technology*, *51*(5), 1450–1484. https://doi.org/10.1111/bjet.12993

Shute, V., & Ventura, M. (2013). Stealth Assessment. In *The SAGE Encyclopedia of Educational Technology* (Issue January 2013, p. 91). SAGE Publications, Inc. https://doi.org/10.4135/9781483346397.n278

Spikol, D., Prieto, L. P., Rodríguez-Triana, M. J., Worsley, M., Ochoa, X., Cukurova, M., Vogel, B., Ruffaldi, E., & Ringtved, U. L. (2017). Current and future multimodal learning analytics data challenges. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 518–519. https://doi.org/10.1145/3027385.3029437

Srivastava, N., Nawaz, S., Newn, J., Lodge, J., Velloso, E., M. Erfani, S., Gasevic, D., & Bailey, J. (2021). Are you with me? Measurement of Learners' Video-Watching Attention with Eye Tracking. *LAK21: 11th International Learning Analytics and Knowledge Conference*, 88–98. https://doi.org/10.1145/3448139.3448148

Stadler, M., Niepel, C., & Greiff, S. (2019). Differentiating between static and complex problems: A theoretical framework and its empirical validation. *Intelligence*, *72*, 1–12. https://doi.org/10.1016/j.intell.2018.11.003

Stark, R., Kopp, V., & Fischer, M. R. (2011). Case-based learning with worked examples in complex domains: Two experimental studies in undergraduate medical education. *Learning and Instruction*, *21*(1), 22–33. https://doi.org/10.1016/j.learninstruc.2009.10.001

Steenbergen-Hu, S., & Cooper, H. (2014). A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. *Journal of Educational Psychology*, *106*(2), 331–347. https://doi.org/10.1037/a0034752

Suárez, N., Regueiro, B., Estévez, I., del Mar Ferradás, M., Guisande, M. A., & Rodríguez, S. (2019). Individual Precursors of Student Homework Behavioral Engagement: The Role of Intrinsic Motivation, Perceived Homework Utility and Homework Attitude. *Frontiers in Psychology*, *10*, 941. https://doi.org/10.3389/fpsyg.2019.00941

Sweller, J. (2010). Element Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load. *Educational Psychology Review*, *22*(2), 123–138. https://doi.org/10.1007/s10648-010-9128-5

Sweller, J. (2024). Cognitive load theory and individual differences. *Learning and Individual Differences*, *110*, 102423. https://doi.org/10.1016/j.lindif.2024.102423

Tabak, I., & Kyza, E. A. (2018). Research on Scaffolding in the Learning Sciences. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Reimann (Eds.), *International Handbook of the Learning Sciences* (1st ed., pp. 191–200). Routledge. https://doi.org/10.4324/9781315617572-19

Taub, M., Azevedo, R., Rajendran, R., Cloude, E. B., Biswas, G., & Price, M. J. (2021). How are students' emotions related to the accuracy of cognitive and metacognitive processes during learning with an intelligent tutoring system? *Learning and Instruction*, *72*, 101200. https://doi.org/10.1016/j.learninstruc.2019.04.001

Taub, M., Mudrick, N. V., Azevedo, R., Millar, G. C., Rowe, J., & Lester, J. (2017). Using multi-channel data with multi-level modeling to assess in-game performance during gameplay with Crystal Island. *Computers in Human Behavior*, *76*, 641–655. https://doi.org/10.1016/j.chb.2017.01.038

Tepgec, M., & Ifenthaler, D. (2022, November 8). Learning Analytics based interventions: A systematic review of experimental studies. *Proceeedings of the 19th International Conference on Cognition and Exploratory Learning in the Digital Age (CELDA 2022)*. 19th International Conference on Cognition and Exploratory Learning in the Digital Age (CELDA 2022). https://doi.org/10.33965/CELDA2022_202207C045

Tetzlaff, L., Schmiedek, F., & Brod, G. (2021). Developing Personalized Education: A Dynamic Framework. *Educational Psychology Review*, *33*(3), 863–882. https://doi.org/10.1007/s10648-020-09570-w

Tobias, S. (1989). Another Look at Research on the Adaptation of Instruction to Students Characteristics. *Educational Psychologist*, *24*(3), 213–227. https://doi.org/10.1207/s15326985ep2403_1

Tsarava, K., Moeller, K., Román-González, M., Golle, J., Leifheit, L., Butz, M. V., & Ninaus, M. (2022). A cognitive definition of computational thinking in primary education. *Computers & Education*, *179*, 104425. https://doi.org/10.1016/j.compedu.2021.104425

Tschisgale, P., Wulff, P., & Kubsch, M. (2023). Integrating artificial intelligence-based methods into qualitative research in physics education research: A case for computational grounded theory. *Physical Review Physics Education Research*, *19*(2), 020123. https://doi.org/10.1103/PhysRevPhysEducRes.19.020123

Valle, N., Antonenko, P., Dawson, K., & Huggins-Manley, A. C. (2021). Staying on target: A systematic literature review on learner-facing learning analytics dashboards. *British Journal of Educational Technology*, *52*(4), 1724–1748. https://doi.org/10.1111/bjet.13089

Van Der Graaf, J., Raković, M., Fan, Y., Lim, L., Singh, S., Bannert, M., Gašević, D., & Molenaar, I. (2023). How to design and evaluate personalized scaffolds for self-regulated learning. *Metacognition and Learning*, *18*(3), 783–810. https://doi.org/10.1007/s11409-023-09361-y

Van Gog, T., & Rummel, N. (2010). Example-Based Learning: Integrating Cognitive and Social-Cognitive Research Perspectives. *Educational Psychology Review*, *22*(2), 155–174. https://doi.org/10.1007/s10648-010-9134-7

Van Otterlo, M., & Wiering, M. (2012). Reinforcement Learning and Markov Decision Processes. In M. Wiering & M. Van Otterlo (Eds.), *Reinforcement Learning* (Vol. 12, pp. 3–42). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-27645-3_1

VanLehn, K. (2011). The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems. *Educational Psychologist*, *46*(4), 197–221. https://doi.org/10.1080/00461520.2011.611369

Vanneste, P., Raes, A., Morton, J., Bombeke, K., Van Acker, B. B., Larmuseau, C., Depaepe, F., & Van Den Noortgate, W. (2021). Towards measuring cognitive load through multimodal physiological data. *Cognition, Technology & Work*, *23*(3), 567–585. https://doi.org/10.1007/s10111-020-00641-0

Verbert, K., Govaerts, S., Duval, E., Santos, J. L., Van Assche, F., Parra, G., & Klerkx, J. (2013). Learning dashboards: An overview and future research opportunities. *Personal and Ubiquitous Computing*. https://doi.org/10.1007/s00779-013-0751-2

Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in Human Behavior*, *89*, 98–110. https://doi.org/10.1016/j.chb.2018.07.027

Vogel, F., Wecker, C., Kollar, I., & Fischer, F. (2017). Socio-Cognitive Scaffolding with Computer-Supported Collaboration Scripts: A Meta-Analysis. *Educational Psychology Review*, *29*(3), 477–511. https://doi.org/10.1007/s10648-016-9361-7

Vuorenmaa, E., Järvelä, S., Dindar, M., & Järvenoja, H. (2023). Sequential Patterns in Social Interaction States for Regulation in Collaborative Learning. *Small Group Research*, *54*(4), 512–550. https://doi.org/10.1177/10464964221137524

Wagenmakers, E.-J., Sarafoglou, A., Aarts, S., Albers, C., Algermissen, J., Bahník, Š., Van Dongen, N., Hoekstra, R., Moreau, D., Van Ravenzwaaij, D., Sluga, A., Stanke, F., Tendeiro, J., & Aczel, B. (2021). Seven steps toward more transparency in statistical practice. *Nature Human Behaviour*, *5*(11), 1473–1480. https://doi.org/10.1038/s41562-021-01211-8

Wiedbusch, M. D. ., Lester, J., & Azevedo, R. (2023). A multi-level growth modeling approach to measuring learner attention with metacognitive pedagogical agents. *Metacognition and Learning*, *18*(2), 465–494. https://doi.org/10.1007/s11409-023-09336-z

Wiedbusch, M. D., Dever, D., Li, S., Amon, M. J., Lajoie, S., & Azevedo, R. (2023). Measuring Multidimensional Facets of SRL Engagement with Multimodal Data. In V. Kovanovic, R. Azevedo, D. C. Gibson, & D. Lfenthaler (Eds.), *Unobtrusive Observations of Learning in Digital Environments* (pp. 141–173). Springer International Publishing. https://doi.org/10.1007/978-3-031-30992-2_10

Wiedbusch, M. D., Kite, V., Yang, X., Park, S., Chi, M., Taub, M., & Azevedo, R. (2021). A Theoretical and Evidence-Based Conceptual Design of MetaDash: An Intelligent Teacher Dashboard to Support Teachers' Decision Making and Students' Self-Regulated Learning. *Frontiers in Education*, *6*, 570229. https://doi.org/10.3389/feduc.2021.570229

Wilcox, R. (2022). *Introduction to Robust Estimation and Hypothesis Testing*. Elsevier.

Winne, P. H. (2017). Cognition and Metacognition within Self-Regulated Learning. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of Self-Regulation of Learning and Performance* (2nd ed., pp. 36–48). Routledge. https://doi.org/10.4324/9781315697048-3

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research. *Frontiers in Psychology*, *10*, 3087. https://doi.org/10.3389/fpsyg.2019.03087

Witte, M., Ninaus, M., Kober, S. E. S. E. S. E., Neuper, C., & Wood, G. (2015). Neuronal Correlates of Cognitive Control during Gaming Revealed by Near-Infrared Spectroscopy. *PLOS ONE*, *10*(8), e0134816. https://doi.org/10.1371/journal.pone.0134816

Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., & Drachsler, H. (2021). Are We There Yet? - A Systematic Literature Review on Chatbots in Education. *Frontiers in Artificial Intelligence*, *4*, 654924. https://doi.org/10.3389/frai.2021.654924

Wong, J., Baars, M., De Koning, B. B., Van Der Zee, T., Davis, D., Khalil, M., Houben, G.-J., & Paas, F. (2019). Educational Theories and Learning Analytics: From Data to Knowledge: The Whole Is Greater Than the Sum of Its Parts. In D. Ifenthaler, D.-K. Mah, & J. Y.-K. Yau (Eds.), *Utilizing Learning Analytics to Support Study Success* (pp. 3–25). Springer International Publishing. https://doi.org/10.1007/978-3-319-64792-0_1

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Wortha, F., Azevedo, R., Taub, M., & Narciss, S. (2019). Multiple Negative Emotions During Learning With Digital Learning Environments – Evidence on Their Detrimental Effect on Learning From Two Methodological Approaches. *Frontiers in Psychology*, *10*, 2678. https://doi.org/10.3389/fpsyg.2019.02678

Xhakaj, F., Aleven, V., & McLaren, B. M. (2017). Effects of a Dashboard for an Intelligent Tutoring System on Teacher Knowledge, Lesson Plans and Class Sessions. In E. André, R. Baker, X. Hu, Ma. M. T. Rodrigo, & B. Du Boulay (Eds.), *Artificial Intelligence in Education* (Vol. 10331, pp. 582–585). Springer International Publishing. https://doi.org/10.1007/978-3-319-61425-0_69

Yan, L., Martinez-Maldonado, R., & Gasevic, D. (2024). Generative Artificial Intelligence in Learning Analytics: Contextualising Opportunities and Challenges through the Learning Analytics Cycle. *Proceedings of the 14th Learning Analytics and Knowledge Conference*, 101–111. https://doi.org/10.1145/3636555.3636856

Zhan, Z., Yang, Q., Luo, L., & Zhang, X. (2023). Applying functional near-infrared spectroscopy (fNIRS) in educational research: A systematic review. *Current Psychology*. https://doi.org/10.1007/s12144-023-05094-y

Zhou, B., Bau, D., Oliva, A., & Torralba, A. (2019). Interpreting Deep Visual Representations via Network Dissection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *41*(9), 2131–2145. https://doi.org/10.1109/TPAMI.2018.2858759

Zhu, M., Lee, H.-S., Wang, T., Liu, O. L., Belur, V., & Pallant, A. (2017). Investigating the impact of automated feedback on students' scientific argumentation. *International Journal of Science Education*, *39*(12), 1648–1668. https://doi.org/10.1080/09500693.2017.1347303

Zhu, M., Liu, O. L., & Lee, H.-S. (2020). The effect of automated feedback on revision behavior and learning gains in formative assessment of scientific argument writing. *Computers & Education*, *143*, 103668. https://doi.org/10.1016/j.compedu.2019.103668

Zhu, M., Sari, A. R., & Lee, M. M. (2022). Trends and Issues in MOOC Learning Analytics Empirical Research: A Systematic Literature Review (2011–2021). *Education and Information Technologies*, *27*(7), 10135–10160. https://doi.org/10.1007/s10639-022-11031-6

Zimmerman, B. J. (2002). Becoming a Self-Regulated Learner: An Overview. *Theory Into Practice*, *41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Highlights

- Introduces the Closed-Loop Learning Analytics Framework
- Emphasizes integration of multimodal data for deeper learning insights
- Reviews advancements in data collection, processing, and personalization
- Emphasizes the impact of AI on adaptive and personalized learning experiences
- Stresses the role of educational and psychological theory for learning analytics

Declaration of interest

The authors have no conflicts of interest to declare.