

## The Sequence Matters in Learning - A Systematic Literature Review

Valle Torre, M.; Oertel, Catharine; Specht, M.M.

**DOI**

[10.1145/3636555.3636880](https://doi.org/10.1145/3636555.3636880)

**Publication date**

2024

**Document Version**

Final published version

**Published in**

LAK 2024 Conference Proceedings - 14th International Conference on Learning Analytics and Knowledge

**Citation (APA)**

Valle Torre, M., Oertel, C., & Specht, M. M. (2024). The Sequence Matters in Learning - A Systematic Literature Review. In *LAK 2024 Conference Proceedings - 14th International Conference on Learning Analytics and Knowledge: Proceedings of the 14th Learning Analytics and Knowledge Conference* (pp. 263–272). (ACM International Conference Proceeding Series). ACM.  
<https://doi.org/10.1145/3636555.3636880>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.



# The Sequence Matters in Learning - A Systematic Literature Review

Manuel Valle Torre  
m.valletorre@tudelft.nl  
Delft Institute of Technology  
Delft, Netherlands

Catharine Oertel  
c.r.m.oertel@tudelft.nl  
Delft Institute of Technology  
Delft, Netherlands

Marcus Specht  
m.m.specht@tudelft.nl  
Delft Institute of Technology  
Delft, Netherlands

## ABSTRACT

Describing and analysing learner behaviour using sequential data and analysis is becoming more and more popular in Learning Analytics. Nevertheless, we found a variety of definitions of learning sequences, as well as choices regarding data aggregation and the methods implemented for analysis. Furthermore, sequences are used to study different educational settings and serve as a base for various interventions. In this literature review, the authors aim to generate an overview of these aspects to describe the current state of using sequence analysis in educational support and learning analytics. The 74 included articles were selected based on the criteria that they conduct empirical research on an educational environment using sequences of learning actions as the main focus of their analysis. The results enable us to highlight different learning tasks where sequences are analysed, identify data mapping strategies for different types of sequence actions, differentiate techniques based on purpose and scope, and identify educational interventions based on the outcomes of sequence analysis.

## CCS CONCEPTS

• **Applied computing** → **Interactive learning environments; Computer-assisted instruction;** • **General and reference** → **Surveys and overviews.**

## KEYWORDS

Learning Analytics, Learning Sequences, Sequence Analysis, Literature Review

### ACM Reference Format:

Manuel Valle Torre, Catharine Oertel, and Marcus Specht. 2024. The Sequence Matters in Learning - A Systematic Literature Review. In *The 14th Learning Analytics and Knowledge Conference (LAK '24)*, March 18–22, 2024, Kyoto, Japan. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3636555.3636880>

## 1 INTRODUCTION

Educational Technologies (EdTech) have rapidly scaled teachers' capacity for instruction, for example through live, virtual and recorded lectures [20, 56]. EdTech, has also increased the opportunity for learners to actively engage with learning materials, from practice

problems in electronic environments to automated assessments, or annotation tools [16, 44]. However, some processes that link teaching and learning in a continuous cycle, such as formative assessment, and personalized support and feedback, have not scaled at the same rate [80]. From its origins, Learning Analytics (LA) intends to close this cycle using data generated by learners in educational systems, and processing it to obtain insights grounded in learning theory and provide (semi) automated learning support and feedback [13].

Existing implementations of LA, such as dashboards or nudges, are often based on descriptive analytics of performance metrics [90], designed for summative assessment and (self) monitoring [97]. However, analytics of the learning process with the detail level to enable timely assessment and support are an ongoing development [80]. Learning is a process that unfolds and changes over time, and aggregating or counting the actions the learner executes is limited in its narrative power [76]. For that reason, it is critical to understand the execution of learning tasks as a temporal and sequential notion [53]. Using sequential data to observe, understand, and support learning is not exclusive to LA, but has a history in fields such as Educational Data Mining (EDM), Intelligent Tutoring Systems (ITS), and Artificial Intelligence in Education (AIED) [93]. This results in a lack of common definitions of concepts such as learning sequences or sequence units. Furthermore, there are no established best practices of the analysis methods that can be applied for specific purposes, educational settings or scenarios [41]. As a consequence, research is often in the exploratory stage, using data from scenarios where the analysed course or task is over, without intervention [55].

In this article, we work towards a broader understanding of the variety of definitions, data points, methods, and interpretations used for the analysis of learning as a sequence. We aim to enable a broader understanding of the learner context to be taken into account for intelligent feedback and analytics interventions. In the following, we will first introduce the state-of-the-art of learning sequence analytics and then present the analysis of 74 articles considering the given dimensions.

## 2 SEQUENCE ANALYSIS AND LEARNING SEQUENCES

Sequence Analysis (SA) of human behaviour, where a sequence is an ordered list of events, allows the contextualization of these events, their relationships with each other, and the effects of part of or the whole sequence [1]. SA is especially useful to research phenomena that occur as a process across time, where the use of static attributes and frequency rates may not be informative enough to understand their development [1]. The analysed sequences can



This work is licensed under a Creative Commons Attribution International 4.0 License.

LAK '24, March 18–22, 2024, Kyoto, Japan  
© 2024 Copyright held by the owner/author(s).  
ACM ISBN 979-8-4007-1618-8/24/03.  
<https://doi.org/10.1145/3636555.3636880>

include actions executed by a person as well as reactions of others in their environment, i.e. a peer or an intelligent agent.

Sequences can have different interpretations in education, in this review, we focus only on the sequences of actions that learners execute during a learning task. For example, the steps followed to solve a problem [93] or the sequences of answers that learners provide in guided practice environments, and how certain sequences may relate to learners' behaviours [91] or characteristics [5].

## 2.1 Sequences of Actions in Learning Tasks

Learning tasks refer to situations when learners actively engage with the learning materials which require them to integrate domain-specific knowledge and procedures with more general critical thinking and analysis skills [62], such as solving complex problems [40] or scientific inquiry tasks [93]. For tasks in different domains, there are behavioural patterns that appear with different levels of practice. For example, experts in a domain may show deeper categorization of a problem, use forward-thinking strategies in their execution of the solution or iteratively evaluate intermediate solutions [62]. In these tasks, learners have a large number of operators or actions available in the problem space, or learning environment, making the combinations leading to a solution difficult to observe. Furthermore, defining and describing such combinations is not trivial, since they are "automated", or stored in one's memory as procedural learning, making it difficult to explain by experienced individuals [62]. As a consequence, the assessment and monitoring of learners are particularly difficult for teachers, especially considering the scale provided by educational technologies and learning platforms [32]. The data produced by these systems, however, can be leveraged to scale the capabilities of teachers, by analysing sequences of learner's data records to identify behaviours and reveal learning processes [51].

## 2.2 Trace Data Sequences and Analysis

The reliability of trace data comes from its ability to express dynamic events, more than as aggregated, static aspects of a measure [93]. Using analytics to find patterns in the sequences that indicate specific tactics or common issues can help practitioners obtain insights into the use of different strategies. Such patterns can be a set of learning actions that occur in succession for a proportion of the learners, from partial segments to complete sequences. In addition, given that the patterns are extracted from sequences of actions, they can be interpreted by the teacher and even learners themselves [86]. For example, in a 3D geometry game where learners can insert, remove, and rotate shapes to solve puzzles, teachers cannot assess nor give feedback to learners based on the gameplay records without assistance. Common patterns that lead to a failed assignment can be identified as mistakes and teachers can prioritize which to address [25].

Analysing sequences of trace data throughout a learning task offers practical advantages as well [87]: when systems process data traces as sequences of learner actions, there is a constant input of information to incrementally update the learner model, enabling the system with the flexibility to constantly adapt to the learner's level or intervene whenever necessary [15]. This allows practitioners, and eventually, an intelligent system, to predict if the current path

is trending towards an undesirable outcome, to determine which action patterns are leading to that result and how to leverage these insights to optimise learning [72, 85].

## 2.3 Related Work

The analysis of learning sequences is emerging from several fields, and while there are no surveys specifically on the topic, there are related works from the EDM, LA and AIED communities. Regarding the use of educational data to observe learner behaviour and predict performance, Xiao et al. [100] analyse EDM methods in the age of big data to discover the trends around the complete data mining pipeline, such as data collection and preprocessing, feature engineering, classifier and ensemble method selection, and interpretation of results. While they mention techniques that analyse sequential data, there are no definitions of learning sequences or sequence units. On the other hand, the survey by Bogarín et al. [7] focuses on Educational Process Mining (EPM), a field that studies the process-oriented view of EDM, particularly as an end-to-end process. The authors describe Virtual Learning Environments (VLE) and the Event Logs they produce as valuable data sources for sequential analysis, as they are detailed records of learners' actions. They mention methods handling learning as sequences, including Sequential Pattern Mining, but the definitions outside of process mining are not addressed.

Regarding the use of trace data to observe learners' actions, the review by Wang et al. [93] analyses research on the use of log data from Open-Ended Learning Environments (OELEs) to understand scientific inquiry and problem-solving. Their conceptual framework consists of 4 points: the competencies measured, mainly problem-solving and inquiry; the observations used, ranging from multiple-choice questions to log data; the feature extraction and modelling processes and the interpretation of the analysis. While the focus is not only on sequential methods, almost half of the 70 reviewed papers use sequential features for analysis, with 29 applying pattern mining to identify behaviours. The review by Nadimpally et al. [58] surveys Artificial Intelligence techniques for adaptive learning systems, identifying two main types: to model learner behaviour and to organize learning content. Concerning modelling learner behaviour, they found that the most common methods are Similarity-based Pattern Recognition, Probabilistic Graphical Models, and Deep Learning: all of them capable of processing sequential data. Existing literature explores the growing use of temporal data to identify, classify or predict learner behaviours, which often involves sequence analysis. However, a focused study of when to use sequence analysis, the data units that can be leveraged and which methods to implement, is less covered.

## 3 RESEARCH QUESTIONS

Addressing the shortcomings from the previous section, this study intends to establish a common ground for Sequence Learning Analytics from the literature, so researchers can use this to find similar works, compare their approach to the state-of-the-art, and leverage insights of existing work, allowing the community to also focus on the intervention design. In this review, a **learning sequence** is defined as an ordered list of context-specific actions, with each

**unit** being a systematically mapped learning action from one or more data traces. The following questions guide our analysis:

- (1) How are learning sequences analyzed to describe and support learners' actions in computer systems?
  - (a) Which types of *learning tasks* are analysed as sequences?
  - (b) What is the motivation or *purpose* when using learning sequences?
  - (c) Which *actions* are used as the learning sequence units?
  - (d) Which *methods* are used to analyse learning sequences? And what are the interpretations of the patterns that result from them?
  - (e) Which *educational interventions* are designed with the obtained insights? Which learning principles are used to support and evaluate them?

## 4 METHOD

In this review, we followed the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) [67], to answer the questions above, in Section 3. The queries, information sources and eligibility criteria are described in this section, followed by the document collection and inclusion process, and the data extraction method.

The queries were executed in SCOPUS and Web of Science (WoS), for the domains of [Educational Data Mining \(EDM\)](#), [Artificial Intelligence in Education \(AIED\)](#), [Learning Analytics \(LA\)](#), [Intelligent Tutors and Learner Modeling](#), exploring the fields of title, abstract and author-defined keywords. SCOPUS and WoS were selected as they include most of the top publications in such domains, such as the proceedings of the conferences on Learning Analytics and Knowledge, Educational Data Mining, and the International Conference on Computers in Education, as well as the Journal of Learning Analytics, British Journal of Educational Technology, Journal of Educational Data Mining, and Computers and Education: Artificial Intelligence. The queries were executed on September 2023, including articles from 2010 to 2023, and they focus on the use of [sequential analysis](#) in an [educational setting](#), where the obtained patterns are leveraged for an [educational intervention](#). Two queries were executed, one per database, using the query below, with the results described in Section 4.2.

*("data mining" OR "artificial intelligence" OR "learning analytics" OR "intelligent tutor\*" OR "learner model\*") AND (sequen\* OR procedur\* OR series ) AND (educatio\* OR student OR teacher OR instructor ) AND (feedback OR assessment OR predict\* )*

### 4.1 Study Selection Process and Eligibility Criteria:

The results were included if they were full-length articles, published in a peer-reviewed source, and written in English. This was followed by two rounds of screening, the first one based on the abstracts and the second on full text for those that were unclear from the first screening. To be included, the publication had to be implemented in an educational setting, where the learning sequences refer to an ordered series of actions during a learning task. Non-empirical works are excluded, as are works that only mention education as part of the institution name or a possible use case. Additionally, publications were excluded if they analysed unrelated sequencing,



**Figure 1: Document Flow Diagram**

like proteins, or where sequences were part of the methodology, such as “a sequence of trials” or “a series of interviews”. Finally, works exploring learner patterns across multiple tasks were excluded, such as navigation patterns on the elements accessed in MOOC or the courses in a curriculum.

### 4.2 Data Collection Process and Data Items

Reference lists were exported from SCOPUS and WoS, and then transferred to Google Sheets for screening. The reason for exclusion or inclusion was recorded for each article, and the values are shown in Figure 1. Of the articles obtained from the queries, 953 were unique, then in the abstract screening 717 were excluded since they were not in an educational setting, not empirical work, the use of learner action sequences was not part of the analysis or did not focus on a single learning task. The full text of the remaining 236 articles was screened, and 162 were excluded based on the criteria described above, resulting in 74 articles included. Document collection and inclusion were performed by the first author, using the criteria previously defined with the rest of the team.

### 4.3 Analysis Procedure

The included studies were coded by the first author using ATLAS.ti (a qualitative research software) with the Research Questions from Section 3) as a guide, following negotiated consensual validation with the rest of the authors to iteratively review the codes [77]. For each question, we determined 3 or 4 possible categories from the related work in Section 2.3, and then the full text of the articles was reviewed and coded using the questions and categories. If a code answered one of the questions but did not fit any of the categories, a new category was created. Additionally, a second coder reviewed a sample of 10% of the articles and the results were compared to ensure the reliability of the coding process, which showed an inter-rater reliability kappa of 0.922, where any differences were discussed and addressed.

For the first question, we identified the learning tasks considering the educational construct they are based on: deliberate practice, solving a problem or scientific inquiry [93]. The first one refers to tasks where learners practice a skill through repetition, usually with a system tracking their progress and adjusting accordingly. The second one refers to tasks where learners have to solve a complex problem in a virtual environment, and the third one refers to tasks where learners have to obtain and aggregate information.

On the other hand, the purpose of sequence analysis can be defined similarly to the goal of data mining and analytics tasks, such as clustering, classification or prediction [100], in line with the findings in [93]. While all the reviewed works identify patterns using sequence analysis, some of them go further and compare the patterns between groups of learners, for example, to identify behaviour differences among learners who received different experimental treatments, or to classify them by the use of strategies. Finally, some works use the identified patterns to predict the next action of a sequence, the success of such action, or a learner's state.

Regarding the units of the sequences, the reviewed research uses data sources such as log files or event records [95], which have to be mapped to representations of meaningful actions, to be analysed and understood as part of the learning sequence [59]. These sequence units can be categorised as learning actions, which are domain or system-specific actions such as selecting and using a system tool in an OELE [93], or metacognitive actions, usually defined based on an educational theory, such as self-explanation [38]. The third category is problem attempts, the series of submissions in a guided practice system [58, 101].

For the implemented methods, we obtained the categories of methods that use sequential data from related literature on AI in Education and Process Mining [7, 58]: Probability Modeling, Deep Learning, Sequential Pattern Mining, Educational Process Mining and Knowledge Tracing. Additionally, we identify the scope of the resulting sequences as the reference for interpretation, from system-specific tactics to general strategies for a complete task [7].

Regarding interventions and educational theories, we seek to identify which works go beyond the observation and exploration stage, which is why we included *assessment, prediction and feedback* in our query, in Section 4. However, considering the challenge of implementing and evaluating an educational intervention [55, 71], we classify if articles implement an intervention, or if they explicitly mention a potential intervention or not. Additionally, the impact of interventions relies on a strong theoretical foundation [96], so we identify the educational theories used to support or interpret the results of sequential analysis [37].

#### 4.4 Limitations

The search was limited to Web of Science and SCOPUS, while these libraries include a wide range of publications, the results are not exhaustive. Additionally, there are no established practices to mention sequence analysis in the title, keywords, or abstract, in consequence, some articles may not have appeared in our queries. Furthermore, abstracts were used to assess the first round of inclusion, there is the possibility that articles were erroneously excluded. Regarding the analysis, the articles were coded only by the first author, however, we addressed this limitation by following negotiated consensual

validation with the author team [77] and assessing the agreement of a sample with a second coder.

## 5 RESULTS

Our analysis included 74 articles, with the number of publications doubling in the second half of our search range, suggesting a constant increase in interest in the field, shown in Figure 2. The articles were analysed following the Research Questions in Section 3, to establish the trends of the different types of learning tasks, the different purposes of the sequence analysis, the types of units and analysis methods used, and any implemented interventions. The following sections address each of these questions, with the results available as an open-access dataset [89], and as an interactive interface at <https://sequence-dashboard.learn.ewi.tudelft.nl/>.

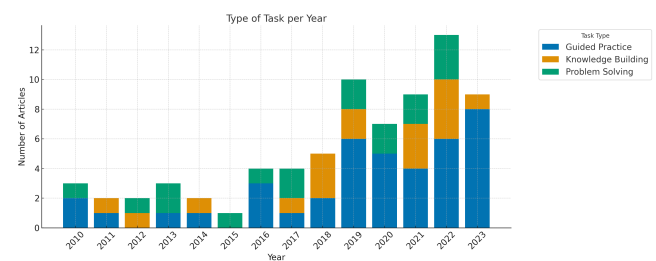


Figure 2: Articles, and Task Types, per Year

### 5.1 Which types of learning tasks are analysed using sequences?

This review focuses on tasks that require active engagement from the learner, which instructional design models such as the 4C/ID model [40], consider as authentic problems and practices. However, they can follow different educational designs. We identify three main types of tasks: guided practice, problem-solving, and knowledge building; the yearly number of publications is shown in Figure 2. The first type is *guided practice*, characterized by a set of practice questions that are part of a topic, where a system uses the sequence of answers to constantly measure and adapt to the learner's knowledge level. We registered 40 articles of this type, 2 of them using dialogue [27, 57], and 2 as a quiz [35, 68].

The second type of task is *problem-solving*, a complex task integrating domain knowledge with domain-specific and general skills. In these tasks, the learner's interactions with the system can be interpreted as the operations executed towards a solution state [4]. While these tasks generally have a defined objective, they can be satisfied with different sequences of actions or applied operators, from a wide variety of actions available within the educational system. The 17 articles using data from a variety of problem-solving tasks, including programming assignments [36], mechanical engineering problems [81] or procedures in a biotechnology virtual lab [72, 73].

Finally, some of the tasks require that the learner collects and processes information, from video lectures or reading materials, and then generates and evaluates a representation of knowledge, such as a concept map [8]. In these tasks, the final result may be similar for learners who followed different strategies, or sequences of steps, and would benefit from personalized feedback. We identified 17

articles that can be classified as *knowledge-building* tasks, for example, where a learner interacts with articles on climate change and then generates a causal model to explain the relationships between the learned concepts [64].

In summary, we identified 40 guided practice tasks, 17 problem-solving and 17 knowledge-building tasks, with the first type gaining ground in recent years, shown in Figure 2. This may be due to the introduction of increasingly sophisticated deep learning methods or the availability of public datasets [61, 98, 101], eliminating the need to design new systems or collect data. Additionally, it might be related to the research objectives, described below.

## 5.2 What is the motivation or *purpose* of using learning sequences for analysis?

There are many affordances to the use of learning sequences and sequence analysis methods in EDM and LA, such as the ability to extract the sequential patterns from data [49], to identify them as tactics or strategies [54], or to include temporal structures in behaviour analysis [44]. Most articles identify the patterns as a learning behaviour, tactic or strategy, however, some of them go further and use these patterns for classification or for prediction. We found 8 cases of *pattern identification*, for example, as the self-regulatory processes that occur before a help-seeking event. These patterns can be used to anticipate the struggle of learners who may not use help-seeking tools correctly [42]. Another case of identification is when groups of learners receive different experimental treatments, such as a feedback strategy [83], and then the patterns of the groups are compared. This means that in these 6 cases, the identified sequences are used as part of the evaluation of an educational intervention [39].

Going a step further, in the 24 articles where the purpose of the sequence is to serve for *classification*, researchers first separate the learners by a given characteristic, such as proficiency in the task [69] and then identify the patterns per group. In these cases, like in traditional data mining classification tasks, a small percentage of the data is used to evaluate the accuracy of the method, or how well can it classify a learner by their behaviour [31, 69, 70]. In only one article the grouping is done by applying clustering algorithms on the sequences of all learners, and then samples of the groups are reviewed and labelled as a certain strategy, such as *Efficient Completion* [54].

Regarding *prediction*, the purpose is to use a sequence as input but there is a variety of outputs: in 8 cases it is the learner's performance, such as success or failure on the task [12, 81], in 2 is the learner's state such as emotional or cognitive states [5, 15] and for 1 it is the next action they might do [72]. For those works focusing on prediction, this is often with the vision of an early warning system or even automated interventions. Finally, in 24 cases, the purpose is to predict the success of a learner's next action, often to determine their knowledge level for a given topic [92].

## 5.3 Which actions are used as the *units* of learning sequences?

In the analysed articles we found a broad variety of data records logged, including clicks on a web interface, sensor data from haptic devices, and text input in command line interfaces. A single case

uses the data as-is, as the authors consider them detailed enough for the sensor data collected from a dentistry virtual reality simulator [70]. In the rest of the publications, we found three different types of combining data records into traces: learning actions, problem attempts, and meta-cognitive actions. We identified 26 articles that use *learning actions* as units, which usually represent the different ways that learners apply operators in *problem-solving tasks*, defined in Section 5.1, such as editing code, executing a command, or using a tool. Mapping the data records into such actions allows researchers, teachers and students to easily analyse and interpret the results of sequence methods. The downside is that this mapping usually requires expertise with the domain and system, to know which actions are possible, and it also limits the possibility to generalize the analysis and results.

In the second case, identified in 37 of the reviewed works, the collected data records are better portrayed as *problem attempts* because they are always delimited by a submission event: when a student submits a response to be evaluated by the system, usually named Intelligent Tutoring System (ITS), and they correspond to those using the *guided task* type. Considering that there is a clear definition of where one unit starts and ends, there is little focus on the process of mapping data records to units.

Finally, for the 10 articles using *meta-cognitive actions*, we found that they introduced a conceptual background from different learning theories, and then use the selected theory to construct the sequence units. Eight of them use a Self-Regulated Learning (SRL) model to map certain records into a meta-cognitive action. From the rest, one maps the data to meta-cognitive states, which they refer to as on-task and off-task states [83], and the other to a variety of meta-cognitive and cognitive behaviours, such as evaluating, help-seeking, and skipping video [11]. Using conceptually mapped actions reduces the granularity and length of the sequences to be analysed since they usually aggregate data records into just a few possible options, compared with the simpler learning actions above. Additionally, this means less variability during analysis and a more direct interpretation.

The sequence units used for analysis are strongly related to the type of learning task, the 26 articles with learning actions tend to be on problem-solving, using specific actions in specific problems, and the 37 using problem attempts are generally on guided practice tasks, focusing more on the accuracy of the learner's model. Finally, the 10 cases of meta-cognitive actions are usually in the knowledge-building tasks, focusing more on self-regulated learning processes.

## 5.4 Which *methods* are used for the analysis of learning sequences?

The methods in the reviewed articles usually take a set of sequences as input, where each one corresponds to learning actions executed by one learner during a task. The methods that we found range from the manual analysis of game logs [65] to deep learning models capable of predicting the performance of the learner [81], with many others among them. In our analysis, we classify the different methods into 5 categories: Sequential Probability Models, Educational Process Mining, Deep Learning, Knowledge Tracing, and Sequential Pattern Mining (SPM). *Sequential Probability Models*

**Table 1: Methods per Category and Scope**

Scope	Method Category	Analysis Method	References
Macro-level Patterns	Educational Process Mining	Fuzzy mining, PetriNet, Process Maps	[35, 60, 69]
	Other Methods	N-gram processing, Network Visualization	[32, 33, 49]
	Sequential Probability Model	Markov Chain Model, Markov Chain Process, Hidden Markov Model, Microsoft Sequence Clustering, Markov Logic Networks, Probabilistic Automata and Hidden Conditional Random Fields	[16, 17, 27, 29, 38, 52, 64, 70, 72, 73, 86]
Meso-level Patterns	Deep Learning Knowledge Tracing	GRU, MLP, CNN, LSTM, SAE PFA, ModPFA, BKT, DKT, DKVMN, AKT, and many more	[12, 15, 20, 22, 56, 68, 81, 104] [6, 14, 19, 21, 23, 26, 28, 30, 34, 47, 48, 50, 61, 74, 75, 78, 88, 92, 98, 99, 101–103]
	Other Methods	Dynamic Time Warping, Model-free bounded optimization, Course-agnostic Student Performance Model	[10, 54, 79]
Micro-level Patterns	Sequential Pattern Mining	SPAM, Differential, Temporal Interestingness of Patterns in Sequences (D-TIPS), Pex-SPAM, Generalized Sequential Pattern Mining, Frequent Item set,	[2, 3, 8, 24, 25, 31, 39, 43, 45, 57, 63, 87, 91, 94]
	Manual Analysis	Manual Labeling, Sequence of Action Tables, State Permutations	[5, 18, 36, 83]
	Other Methods	Visualization, Lag Sequential Analysis, GSEQ (Generalized Sequential Querier), Time-embedded n-gram	[11, 44, 46, 84]

estimate the probabilities of transition between elements in a sequence, based on the previous one or more elements [17]. On the other hand, *Educational Process Mining* refers to the EDM technique to discover, visualise and compare process models, focusing on the complete sequence [69]. *Deep Learning* methods use Artificial Neural Networks (ANN) to predict the outcome of a sequence, mainly Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks, as they are better suited to handle sequential data than other ANNs [12]. While *Knowledge Tracing* approaches leverage probabilistic, machine learning or deep learning techniques, their main focus is to model the knowledge evolution of a learner throughout their interaction with an educational system [28]. *Sequential Pattern Mining* methods refer to a group of data mining techniques used to find statistically significant patterns—or subsequences—in sequence datasets, such as actions that happen frequently in the same order across learners' records [39]. *Manual Analysis* methods are often used in combination with other methods but are also used on their own in a few cases, for instance by reviewing and classifying the actions of individual learners in a small group [5]. These methods often relate to the scope of the analysis, from the full-task strategies shown by a learner population to short tactics that certain students apply depending on context: the Macro-level Patterns identify strategies that span across the complete task and are analysed for a population or group. The Meso-level refers to behaviours that appear in large part of the task, but the focus is on the individual level. Finally, the Micro-level patterns are short tactics that occur in a particular context. Shown in Table 1, we group the examples of each method and the scope that they belong to.

## 5.5 Educational Interventions - What is the final intervention designed using the obtained insights?

Educational interventions provide students with the support required to acquire the skills and knowledge being taught. However, interventions remain one of the main challenges in the learning analytics cycle [71]. This is reflected in our findings, where most reviewed publications are exploratory studies with potential interventions, for instance, as an early warning system that can notify teachers about learners using tactics that lead to poor performance [94], or automated systems that can detect if a learner might go into a blocking state [15]. There were also 5 prototype teacher dashboards [9, 24, 25, 35, 84], with one of them receiving a positive response from a usability survey with 56 participants [33].

In our review, we found 2 articles that design and implement an intervention. In the first, the authors integrate the patterns obtained to provide suggestions to students in a programming tutoring system. The suggestions included reading the manual or asking a peer for help when an error state is identified, to nudge the learners into a self-regulation strategy, for example, *reflection-then-success* after a sequence that could lead to failure [91]. The impact of this intervention, however, is not evaluated. The second one implemented and evaluated the use of learning sequences, by creating a “good-learner path” model using a Markov Decision Process in a guided practice system, so it can automatically generate the content sequence for first-time users. The model was then evaluated by comparing the test results of learners who followed the created path and learners who were provided with exercises one by one based on their prior results. Their approach produced 19.5% more good learners and 20.5% lower standard deviation than the control group [52].

## 6 DISCUSSION

### 6.1 Learning Tasks and Analysis Purpose

Regarding the *purpose* of sequence analysis, the common objective is to extract patterns from the data and identify the real-world behaviours they describe. In some cases, these patterns are then used to classify the learner between groups or to predict the learner's state, their subsequent action, or the outcome of the task. In most research using problem-solving tasks, patterns are closely related to the scenario, for example, how novices implement functions in programming. A few cases use these scenario-specific patterns to understand more general skills, for instance, identifying if learners follow linear or iterative processes for problem-solving. Similarly, for the knowledge-building tasks, the focus is to find information processing strategies and the meta-cognitive processes executed by the learners [38, 39]. Using an educational lens to determine the purpose of a study, such as complex problem-solving or self-regulation, facilitates the generalisation of the methods and results, regardless of the task.

On the other hand, the purpose of many guided practice tasks is to improve the prediction of the learners' performance, often to evaluate a proposed method against the state-of-the-art, as is the case for most of the 16 articles using the ASSISTments<sup>1</sup> dataset. In these studies, the sequential patterns receive less attention, and educational theory is leveraged in the prediction function, such as learning and forgetting, or while calculating the difficulty of a learning component. While this field has seen important growth with the popularity of Neural Networks (see figure 2), the real-world application and explainability of modern techniques are still key challenges [82].

### 6.2 Learning Actions and Methods

From our analysis, there seem to be increasing similarities in how researchers of different domains use sequences for the analysis of learner behaviour. For example, we found that 25 out of the 37 articles using Problem Attempts as the units of sequence, also integrate contextual information about the learner and their actions before submitting a problem, such as temporal features or their previous knowledge. Another common case is the use of statistical methods in combination with sequential methods, such as using logistic regression with features extracted from a sequential technique. Furthermore, researchers often implement more than one sequence analysis technique, using one to find case-specific action units, and then another to explore how these actions are sequenced in a full task strategy [86]. This may be a consequence of recent efforts towards the transparency and explainability of data-driven analysis, as well as the increasing accessibility of technology and analysis techniques.

The use of machine learning and deep learning methods to power automated or semi-automated decisions has recently started to see widespread implementation. This may be due to the recent efforts to mitigate risks associated with AI, for instance by pursuing transparency, causality, privacy, fairness, trust, usability, and reliability; some of the goals of Explainable AI (XAI) [48]. Explainability is especially important when applied to education, as the explanations can also have an educational component for learners, and it can aid teachers in their assessment of learners and their courses

<sup>1</sup><https://sites.google.com/site/assistmentsdata/datasets>

[48]. Regarding accessibility, different techniques are available via tools and libraries, we found several articles that use the Waikato Environment for Knowledge Analysis (Weka<sup>2</sup>) for multiple machine learning classifiers and predictors, [8, 31, 57, 65, 66, 73, 74], 2 more that apply Sequential Pattern Mining with TraMineR for R<sup>3</sup> [45, 88], as well as Markov Models via MeTA Toolkit<sup>4</sup> [85], and the Microsoft Sequence Clustering Algorithm [73]. It is important to note that these tools were explicitly mentioned in their respective articles, further facilitating reproducibility and comparison. This allows researchers to work with known methods throughout their analysis process, without having to develop, test, and evaluate them from scratch. Finally, we provide the dataset [89] of the reviewed works as an interactive interface<sup>5</sup> where researchers can filter the methods depending on the scope of their study, as shown in Table 1.

### 6.3 Interventions

Some of the reviewed works mention potential interventions that can be designed using the obtained patterns. For example, adaptive scaffolding in problem-solving tasks, personalized feedback for a knowledge-building system, or detection of the learners' level in guided practice. However, only 2 implemented an intervention, with 1 of them evaluating it, while 20 articles had no mention of a potential intervention at all. This is in line with recent publications such as "LAK of Direction", where they found that 11% of their reviewed articles introduced an intervention into a learning environment [55].

The use of learning principles, such as cognitive load [57] or learning and forgetting curves [21], is often present when describing or framing the problem. Very few use learning principles in the interpretation of patterns or interventions. This may be due to the focus on task- or system-specific tactics more than general learning behaviours. However, actionable results in LA research require a strong theoretical foundation, present in the analysis and the interpretation of the findings [96]. The theories mentioned in the framing of the problem should remain through the complete process, such as the use of Neo-Piagetian theory to classify patterns as programming strategies related to expertise [36]. This was the case for the 10 articles where the purpose is to identify meta-cognitive strategies, with 8 of them using a Self-Regulated Learning model to map the data records into activities, interpret the patterns after analysis, and describe a potential intervention [37].

## 7 CONCLUSION

In this document, we reviewed the different learning tasks, research purposes, action units, and methods of learning sequence analysis in educational systems, as well as the interventions that researchers design and propose with the obtained insights. We described the major trends that can help to establish an overview of the use of Learning Sequence Analysis, aggregated in an open repository [89] and an interactive interface available in <https://sequence-dashboard.learn.ewi.tudelft.nl/>. This is to assist researchers in finding, comparing, and using other's efforts as a foundation and move

<sup>2</sup><https://www.cs.waikato.ac.nz/ml/weka/>

<sup>3</sup><http://traminer.unige.ch/>

<sup>4</sup><https://github.com/meta-toolkit/meta>

<sup>5</sup><https://sequence-dashboard.learn.ewi.tudelft.nl/>



past the exploratory stage; as noted in one of the reviewed articles [72], "the community is missing more generally applicable results."

## ACKNOWLEDGMENTS

This project has been funded by the Leiden-Delft-Erasmus Centre for Education and Learning (LDE-CEL). The authors thank the second coder, Rebecca Glans.

## REFERENCES

- Andrew Abbott. 1995. Sequence Analysis: New Methods for Old Ideas. *Annual Review of Sociology* 21 (1995), 93–113. [jstor:2083405](https://doi.org/10.1146/annurev.soc.21.1.93)
- Kamil Akhuseyinoglu and Peter Brusilovsky. 2021. Data-Driven Modeling of Learners' Individual Differences for Predicting Engagement and Success in Online Learning. In *Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '21)*. Association for Computing Machinery, New York, NY, USA, 201–212. <https://doi.org/10.1145/3450613.3456834>
- Kamil Akhuseyinoglu and Peter Brusilovsky. 2022. Exploring Behavioral Patterns for Data-Driven Modeling of Learners' Individual Differences. *Frontiers in Artificial Intelligence* 5 (Feb. 2022), 807320. <https://doi.org/10.3389/frai.2022.807320>
- John R. Anderson. 2000. *Cognitive Psychology and Its Implications, 5th Ed.* Worth Publishers, New York, NY, USA, xvi, 531 pages.
- Ryan S.J.d. Baker, Antonija Mitrović, and Moffat Mathews. 2010. Detecting Gaming the System in Constraint-Based Tutors. *Lecture Notes in Computer Science* 6075 LNCS (2010), 267–278. [https://doi.org/10.1007/978-3-642-13470-8\\_25](https://doi.org/10.1007/978-3-642-13470-8_25)
- Samara Banno, Hayder Murad, and Muntadher Sallal. 2020. Automated Cognitive Analyses for Intelligent Tutoring Systems. In *2020 IEEE/ACM International Conference on Big Data Computing, Applications and Technologies (BDCAT)*. Leicester, UK, 171–178. <https://doi.org/10.1109/BDCAT50828.2020.00007>
- Alejandro Bogarin, Rebeca Cerezo, and Cristóbal Romero. 2018. A Survey on Educational Process Mining. *WIRES Data Mining and Knowledge Discovery* 8, 1 (2018), e1230. <https://doi.org/10.1002/widm.1230>
- Francois Bouchet, Roger Azevedo, John S. Kinnebrew, and Gautam Biswas. 2012. Identifying Students' Characteristic Learning Behaviors in an Intelligent Tutoring System Fostering Self-Regulated Learning. In *International Educational Data Mining Society*. International Educational Data Mining Society, Chania, Greece.
- Sam Bryczynski, Roy P. Pargas, Melanie M. Cooper, Michael Klymkowsky, and Brian C. Dean. 2013. Teaching Data Structures with beSocratic. In *Proceedings of the 18th ACM Conference on Innovation and Technology in Computer Science Education (ITICSE '13)*. Association for Computing Machinery, New York, NY, USA, 105–110. <https://doi.org/10.1145/2462476.2465583>
- Binglin Chen, Matthew West, and Craig Zilles. 2018. Towards a Model-Free Estimate of the Limits to Student Modeling Accuracy. In *Proceedings of the 11th International Conference on Educational Data Mining*, Vol. 11. Raleigh, NC, USA.
- Hsi-Min Chen, Bao-An Nguyen, Chyi-Ren Dow, Nien-Lin Hsueh, and An-Chi Liu. 2022. Exploring Time-Related Micro-Behavioral Patterns in a Python Programming Online Course. *Journal of Information Science & Engineering* 38, 6 (2022).
- Yun-Wei Chu, Elizabeth Tenorio, Laura Cruz, Kerrie Douglas, Andrew S. Lan, and Christopher G. Brinton. 2021. Click-Based Student Performance Prediction: A Clustering Guided Meta-Learning Approach. In *2021 IEEE International Conference on Big Data (Big Data)*. Orlando, FL, USA, 1389–1398. <https://doi.org/10.1109/BigData52589.2021.9671729>
- Doug Clow. 2012. The Learning Analytics Cycle: Closing the Loop Effectively. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK '12)*. Association for Computing Machinery, New York, NY, USA, 134–138. <https://doi.org/10.1145/2330601.2330636>
- Joshua Cook, Collin F. Lynch, Andrew G. Hicks, and Behrooz Mostafavi. 2017. Task and Timing: Separating Procedural and Tactical Knowledge in Student Models. In *Proceedings of the 10th International Conference on Educational Data Mining*. International Educational Data Mining Society, Wuhan, China.
- Jean-Philippe Corbeil, Michel Gagnon, and Philippe R. Richard. 2020. Probabilistic Approaches to Detect Blocking States in Intelligent Tutoring System. In *Intelligent Tutoring Systems: 16th International Conference, ITS 2020, Athens, Greece, June 8–12, 2020, Proceedings*. Springer-Verlag, Berlin, Heidelberg, 79–88. [https://doi.org/10.1007/978-3-030-49663-0\\_11](https://doi.org/10.1007/978-3-030-49663-0_11)
- Daryn A. Dever, Nathan A. Sonnenfeld, Megan D. Wiedbusch, and Roger Azevedo. 2022. Pedagogical Agent Support and Its Relationship to Learners' Self-Regulated Learning Strategy Use with an Intelligent Tutoring System. In *Proceedings of Artificial Intelligence in Education: 23rd International Conference, AIED (Durham, United Kingdom)*. Springer-Verlag, 332–343. [https://doi.org/10.1007/978-3-031-11644-5\\_27](https://doi.org/10.1007/978-3-031-11644-5_27)
- Daryn A. Dever, Nathan A. Sonnenfeld, Megan D. Wiedbusch, S. Grace Schmorow, Mary Jean Amon, and Roger Azevedo. 2023. A Complex Systems Approach to Analyzing Pedagogical Agents' Scaffolding of Self-Regulated Learning within an Intelligent Tutoring System. *Metacognition and Learning* (May 2023). <https://doi.org/10.1007/s11409-023-09346-x>
- H.D. Duong, L. Zhu, Y. Wang, and N.T. Heffernan. 2013. A Prediction Model Uses the Sequence of Attempts and Hints to Better Predict Knowledge: Better to Attempt the Problem First, Rather than Ask for a Hint. In *Proceedings of the 6th International Conference on Educational Data Mining, EDM*.
- Michael Eagle, Albert Corbett, John Stamper, and Bruce McLaren. 2018. Predicting Individualized Learner Models Across Tutor Lessons. In *Proceedings of the 11th International Conference on Educational Data Mining*, Vol. 11. Raleigh, NC, USA.
- H. El Aouifi, M. El Hajji, Y. Es-Saady, and H. Douzi. 2021. Predicting Learner's Performance through Video Sequences Viewing Behavior Analysis Using Educational Data-Mining. *Education and Information Technologies* 26, 5 (2021), 5799–5814. <https://doi.org/10.1007/s10639-021-10512-4>
- W. Gan, Y. Sun, X. Peng, and Y. Sun. 2020. Modeling Learner's Dynamic Knowledge Construction Procedure and Cognitive Item Difficulty for Knowledge Tracing. *Applied Intelligence* 50, 11 (2020), 3894–3912. <https://doi.org/10.1007/s10489-020-01756-7>
- Michael Geden, Andrew Emerson, Jonathan Rowe, Roger Azevedo, and James Lester. 2020. Predictive Student Modeling in Educational Games with Multi-Task Learning. *Proceedings of the AAAI Conference on Artificial Intelligence* 34, 01 (April 2020), 654–661. <https://doi.org/10.1609/aaai.v34i01.5406>
- Jing Geng, Huali Yang, and Shengze Hu. 2023. Noise-Filtering Enhanced Deep Cognitive Diagnosis Model for Latent Skill Discovering. *Intelligent Automation & Soft Computing* 37, 2 (2023), 1311–1324. <https://doi.org/10.32604/iasc.2023.038481>
- Manuel J. Gomez, José A. Ruipérez-Valiente, Pedro A. Martínez, and Yoon Jeon Kim. 2020. Exploring the Affordances of Sequence Mining in Educational Games. In *Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality*. ACM, Salamanca Spain, 648–654. <https://doi.org/10.1145/3434780.3436562>
- Manuel J. Gomez, José A. Ruipérez-Valiente, Pedro A. Martínez, and Yoon Jeon Kim. 2021. Applying Learning Analytics to Detect Sequences of Actions and Common Errors in a Geometry Game. *Sensors* 21, 4 (Jan. 2021), 1025. <https://doi.org/10.3390/s21041025>
- Yue Gong, Joseph E. Beck, and Neil T. Heffernan. 2010. Comparing Knowledge Tracing and Performance Factor Analysis by Using Multiple Model Fitting Procedures. In *Intelligent Tutoring Systems (Lecture Notes in Computer Science)*, Vincent Allevin, Judy Kay, and Jack Mostow (Eds.). Springer, Berlin, Heidelberg, 35–44. [https://doi.org/10.1007/978-3-642-13388-6\\_8](https://doi.org/10.1007/978-3-642-13388-6_8)
- J. González-Brenes, W. Duan, and J. Mostow. 2011. How to Classify Tutorial Dialogue? Comparing Feature Vectors vs. Sequences. In *EDM 2011 - Proceedings of the 4th International Conference on Educational Data Mining*. 169–178. [https://doi.org/10.1007/978-3-642-13388-6\\_8](https://doi.org/10.1007/978-3-642-13388-6_8)
- Yi Guo, Tingnian He, Aibin Li, Zhuoran Li, Yixuan Rong, and Guoqi Liu. 2023. Deep Knowledge Tracking Method Based on DKVTMN-DTCN Model. In *Computer Science and Education (Communications in Computer and Information Science)*, Wenxing Hong and Yang Weng (Eds.). Springer Nature, Singapore, 623–636. [https://doi.org/10.1007/978-981-99-2446-2\\_57](https://doi.org/10.1007/978-981-99-2446-2_57)
- Eun Y. Ha, Jonathan P. Rowe, Bradford W. Mott, and James C. Lester. 2011. Goal Recognition with Markov Logic Networks for Player-Adaptive Games. In *Proceedings of the Seventh AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE '11)*. AAAI Press, Stanford, California, USA, 32–39.
- Yu He, Xinying Hu, Zhongtian Xu, and Guangzhong Sun. 2020. KT-XL: A Knowledge Tracing Model for Predicting Learning Performance Based on Transformer-XL. In *Proceedings of the ACM Turing Celebration Conference - China (ACM TURC '20)*. Association for Computing Machinery, New York, NY, USA, 175–179. <https://doi.org/10.1145/3393527.3393557>
- J. Herold, A. Zundel, and T.F. Stahovich. 2013. Mining Meaningful Patterns from Students' Handwritten Coursework. In *Proceedings of the 6th International Conference on Educational Data Mining, EDM 2013*.
- Nathan Hurtig, Joseph Hollingsworth, Sarah Blankenship, Eileen Kraemer, Murali Sitaraman, and Jason O. Hallstrom. 2022. Network Visualization and Assessment of Student Reasoning About Conditionals. In *Proceedings of the 27th ACM Conference on Innovation and Technology in Computer Science Education Vol. 1*. ACM, Dublin Ireland, 255–261. <https://doi.org/10.1145/3502718.3524793>
- Nathan Hurtig, Joseph Hollingsworth, and Olga Scrivner. 2022. Visualization of Students' Solutions as a Sequential Network. In *2022 IEEE Global Engineering Education Conference (EDUCON)*. 1189–1194. <https://doi.org/10.1109/EDUCON52537.2022.9766502>
- Xuyang Jiang, Yuanxin Ouyang, Zhuang Liu, Wenge Rong, and Zhang Xiong. 2022. MAKT: Multichannel Attention Networks Based Knowledge Tracing with Representation Learning. In *2022 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE)*. 292–299. <https://doi.org/10.1109/TALE52537.2022.9766502>

- 1109/TALE54877.2022.00055
- [35] L. Juháňák, J. Zounek, and L. Rohliková. 2019. Using Process Mining to Analyze Students' Quiz-Taking Behavior Patterns in a Learning Management System. *Computers in Human Behavior* 92 (2019), 496–506. <https://doi.org/10.1016/j.chb.2017.12.015>
- [36] M. Kesselbacher and A. Bollin. 2019. Discriminating Programming Strategies in Scratch: Making the Difference between Novice and Experienced Programmers. In *Proceedings of the 14th Workshop in Primary and Secondary Computing Education*. <https://doi.org/10.1145/3361721.3361727>
- [37] Mohammad Khalil, Paul Prinsloo, and Sharon Slade. 2022. The Use and Application of Learning Theory in Learning Analytics: A Scoping Review. *Journal of Computing in Higher Education* (Oct. 2022). <https://doi.org/10.1007/s12528-022-09340-3>
- [38] J.S. Kinnebrew and G. Biswas. 2011. Modeling and Measuring Self-Regulated Learning in Teachable Agent Environments. *Journal of E-Learning and Knowledge Society* 7, 2 (2011), 19–35.
- [39] J.S. Kinnebrew, D.L.C. Mack, G. Biswas, and C.-K. Chang. 2014. A Differential Approach for Identifying Important Student Learning Behavior Patterns with Evolving Usage over Time. *Lecture Notes in Computer Science* 8643 (2014), 281–292. [https://doi.org/10.1007/978-3-319-13186-3\\_27](https://doi.org/10.1007/978-3-319-13186-3_27)
- [40] Paul Kirschner and Jeroen J. G. van Merriënboer. 2008. *Ten Steps to Complex Learning: A New Approach to Instruction and Instructional Design*. SAGE Publications, Inc., 2455 Teller Road, Thousand Oaks California 91320 United States, 1–244–1–253. <https://doi.org/10.4135/9781412964012.n26>
- [41] Simon Knight, Alyssa Friend Wise, and Bodong Chen. 2017. Time for Change: Why Learning Analytics Needs Temporal Analysis. *Journal of Learning Analytics* 4, 3 (Dec. 2017). <https://doi.org/10.18608/jla.2017.43.2>
- [42] S.P. Lajoie, E.G. Poitras, T. Doleck, and A. Jarrell. 2015. Modeling Metacognitive Activities in Medical Problem-Solving with Bio World. *Intelligent Systems Reference Library* 76 (2015), 323–343. [https://doi.org/10.1007/978-3-319-11062-2\\_13](https://doi.org/10.1007/978-3-319-11062-2_13)
- [43] V. Latypova. 2022. Work with Free Response Implementation Process Analysis Based on Sequential Pattern Mining in Engineering Education. In *2022 VI International Conference on Information Technologies in Engineering Education (Infornio)*, 1–4. <https://doi.org/10.1109/Infornio53888.2022.9782969>
- [44] Shan Li, Juan Zheng, and Susanne Lajoie. 2022. Temporal Structures and Sequential Patterns of Self-Regulated Learning Behaviors in Problem Solving with an Intelligent Tutoring System. *Educational Technology & Society* 25 (March 2022), 1–14. <https://eric.ed.gov/?id=EJ1366535>
- [45] Lin, J., Singh, S., Sha, L., Tan, W., Lang, D., Gasevic, D., and Chen, G. 2022. Is it a good move? Mining effective tutoring strategies from human-human tutorial dialogues. *Future Generation Computer Systems* 127 (2022), 194–207. <https://doi.org/10.1016/j.future.2021.09.001>
- [46] M.-C. Liu, C.-H. Yu, J. Wu, A.-C. Liu, and H.-M. Chen. 2018. Applying Learning Analytics to Deconstruct User Engagement by Using Log Data of MOOCs. *Journal of Information Science and Engineering* 34, 5 (2018), 1175–1186. [https://doi.org/10.6688/JISE.201809\\_34\(5\).0004](https://doi.org/10.6688/JISE.201809_34(5).0004)
- [47] Zitao Liu, Qiongqiong Liu, Jiahao Chen, Shuyan Huang, Boyu Gao, Weiqi Luo, and Jian Weng. 2023. Enhancing Deep Knowledge Tracing with Auxiliary Tasks. In *Proceedings of the ACM Web Conference 2023 (WWW '23)*. Association for Computing Machinery, New York, NY, USA, 4178–4187. <https://doi.org/10.1145/3543507.3583866>
- [48] Yu Lu, Deliang Wang, Penghe Chen, Qinggang Meng, and Shengquan Yu. 2023. Interpreting Deep Learning Models for Knowledge Tracing. *International Journal of Artificial Intelligence in Education* 33, 3 (Sept. 2023), 519–542. <https://doi.org/10.1007/s40593-022-00297-z>
- [49] Erik Lundgren. 2022. Latent Program Modeling: Inferring Latent Problem-Solving Strategies from a PISA Problem-Solving Task. *Journal of Educational Data Mining* 14, 1 (June 2022), 46–80. <https://doi.org/10.5281/zenodo.6686443>
- [50] Joel Mandlazi, Ashwini Jadhav, and Ritesh Ajoodha. 2021. Evaluating Deep Sequential Knowledge Tracing Models for Predicting Student Performance. In *2021 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*, 1–6. <https://doi.org/10.1109/CSDE53843.2021.9718405>
- [51] Wannisa Matcha, Dragan Gašević, Nora'ayu Ahmad Uzir, Jelena Jovanović, Abelardo Pardo, Jorge Maldonado-Mahauad, and Mar Pérez-Sanagustín. 2019. Detection of Learning Strategies: A Comparison of Process, Sequence and Network Analytic Approaches. In *Transforming Learning with Meaningful Technologies (Lecture Notes in Computer Science)*, Maren Scheffel, Julien Broisin, Viktoria Pammer-Schindler, Andri Ioannou, and Jan Schneider (Eds.). Springer International Publishing, Cham, 525–540. [https://doi.org/10.1007/978-3-030-29736-7\\_39](https://doi.org/10.1007/978-3-030-29736-7_39)
- [52] Selly Meliana and Dade Nurjanah. 2018. Adopting Good-Learners' Paths in an Intelligent Tutoring System. In *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, 877–882. <https://doi.org/10.1109/TALE.2018.8615176>
- [53] Inge Molenaar and Alyssa Friend Wise. 2022. Temporal Aspects of Learning Analytics - Grounding Analyses in Concepts of Time. In *The Handbook of Learning Analytics* (2 ed.). SoLAR, Vancouver, Canada, 66–76.
- [54] Fahmid Morshed Fahid, Xiaoyi Tian, Andrew Emerson, Joseph B. Wiggins, Dolly Bounajim, Andy Smith, Eric Wiebe, Bradford Mott, Kristy Elizabeth Boyer, and James Lester. 2021. Progression Trajectory-Based Student Modeling for Novice Block-Based Programming. In *Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*. Association for Computing Machinery, New York, NY, USA, 189–200. <https://doi.org/10.1145/3450613.3456833>
- [55] Benjamin A. Motz, Yoav Bergner, Christopher A. Brooks, Anna Gladden, Geraldine Gray, Charles Lang, Warren Li, Fernando Marmolejo-Ramos, and Joshua D. Quick. 2023. A LAK of Direction: Misalignment between the Goals of Learning Analytics and Its Research Scholarship. *Journal of Learning Analytics* (March 2023), 1–13. <https://doi.org/10.18608/jla.2023.7913>
- [56] Ahmed Ali Mubarak, Han Cao, and Salah A.M. Ahmed. 2021. Predictive Learning Analytics Using Deep Learning Model in MOOCs' Courses Videos. *Education and Information Technologies* 26, 1 (Jan. 2021), 371–392. <https://doi.org/10.1007/s10639-020-10273-6>
- [57] M.H. Myers. 2021. Automatic Detection of a Student's Affective States for Intelligent Teaching Systems. *Brain Sciences* 11, 3 (2021), 1–15. <https://doi.org/10.3390/brainsci11030331>
- [58] Vamsi Krishna Nadimpalli, Florian Hauser, Dominik Bittner, Lisa Grabinger, Susanne Staufer, and Jürgen Mottok. 2023. Systematic Literature Review for the Use of AI Based Techniques in Adaptive Learning Management Systems. In *Proceedings of the 5th European Conference on Software Engineering Education*. ACM, Seon/Bavaria Germany, 83–92. <https://doi.org/10.1145/3593663.3593681>
- [59] John Nesbit, Mingming Zhou, Yabo Xu, and Philip Winne. 2007. Advancing Log Analysis of Student Interactions with Cognitive Tools. In *2th Biennial Conference of the European Association for Research on Learning and Instruction (EARLI)*, 2–20.
- [60] Andres Neyem, Juan Diaz-Mosquera, Jorge Munoz-Gama, and Jaime Navon. 2017. Understanding Student Interactions in Capstone Courses to Improve Learning Experiences. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*. ACM, Seattle Washington USA, 423–428. <https://doi.org/10.1145/3017680.3017716>
- [61] Qin Ni, Tingjiang Wei, Jiabao Zhao, Liang He, and Chanjin Zheng. 2023. HHSKT: A Learner-Question Interactions Based Heterogeneous Graph Neural Network Model for Knowledge Tracing. *Expert Systems with Applications* 215 (April 2023), 119334. <https://doi.org/10.1016/j.eswa.2022.119334>
- [62] Timothy J. Nokes, Christian D. Schunn, and Micki T. H. Chi. 2010. Problem Solving and Human Expertise. In *International Encyclopedia of Education (Third Edition)*, Penelope Peterson, Eva Baker, and Barry McGaw (Eds.). Elsevier, Oxford, 265–272. <https://doi.org/10.1016/B978-0-08-044894-7.00486-3>
- [63] Y.-C. Norm Lien, W.-J. Wu, and Y.-L. Lu. 2020. How Well Do Teachers Predict Students' Actions in Solving an Ill-Defined Problem in STEM Education: A Solution Using Sequential Pattern Mining. *IEEE access : practical innovations, open solutions* 8 (2020), 134976–134986. <https://doi.org/10.1109/ACCESS.2020.3010168>
- [64] Y. Omae, T. Furuya, K. Mizukoshi, T. Oshima, N. Sakakibara, Y. Mizuochi, K. Yatsushiro, and H. Takahashi. 2018. Data Mining for Discovering Effective Time-Series Transition of Learning Strategies on Mutual Viewing-Based Learning. *Journal of Advanced Computational Intelligence and Intelligent Informatics* 22, 7 (2018), 1046–1055. <https://doi.org/10.20965/jaciii.2018.p1046>
- [65] E.V. Owen, G. Anton, and R. Baker. 2016. Modeling User Exploration and Boundary Testing in Digital Learning Games. In *UMAP 2016 - Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*. 301–302. <https://doi.org/10.1145/2930238.2930271>
- [66] Gustavo Padrón-Rivera and Genaro Rebollo-Mendez. 2015. Identifying Affective Trajectories in Relation to Learning Gains During the Interaction with a Tutoring System. In *Artificial Intelligence in Education (Lecture Notes in Computer Science)*, Cristina Conati, Neil Heffernan, Antonija Mitrovic, and M. Felisa Verdejo (Eds.). Springer International Publishing, Cham, 756–759. [https://doi.org/10.1007/978-3-319-19773-9\\_109](https://doi.org/10.1007/978-3-319-19773-9_109)
- [67] Matthew J. Page, Joanne E. McKenzie, Patrick M. Bossuyt, Isabelle Boutron, Tammy C. Hoffmann, Cynthia D. Mulrow, Larissa Shamseer, Jennifer M. Tetzlaff, Elie A. Akl, Sue E. Brennan, Roger Chou, Julie Glanville, Jeremy M. Grimshaw, Asbjørn Hróbjartsson, Manoj M. Lalu, Tianjing Li, Elizabeth W. Loder, Evan Mayo-Wilson, Steve McDonald, Luke A. McGuinness, Lesley A. Stewart, James Thomas, Andrea C. Tricco, Vivian A. Welch, Penny Whiting, and David Moher. 2021. The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *Systematic Reviews* 10, 1 (March 2021), 89. <https://doi.org/10.1186/s13643-021-01626-4>
- [68] S. Qu, K. Li, B. Wu, S. Zhang, and Y. Wang. 2019. Predicting Student Achievement Based on Temporal Learning Behavior in MOOCs. *Applied Sciences (Switzerland)* 9, 24 (2019). <https://doi.org/10.3390/app9245539>
- [69] R. Rajendran, A. Munshi, M. Emara, and G. Biswas. 2018. A Temporal Model of Learner Behaviors in OELEs Using Process Mining. In *ICCE 2018 - 26th International Conference on Computers in Education, Main Conference Proceedings*. 276–285. <https://apscet.net/icce/icce2018/wp-content/uploads/2018/12/C3-07.pdf>

- [70] P. Rhiemora, P. Haddawy, P. Khanal, S. Suebnukarn, and M. N. Dailey. 2010. A Virtual Reality Simulator for Teaching and Evaluating Dental Procedures. *Methods of Information in Medicine* 49, 4 (2010), 396–405. <https://doi.org/10.3414/ME9310>
- [71] Bart Rienties, Simon Cross, and Zdenek Zdrahal. 2017. Implementing a Learning Analytics Intervention and Evaluation Framework: What Works? In *Big Data and Learning Analytics in Higher Education: Current Theory and Practice*, Ben Kei Daniel (Ed.). Springer International Publishing, Cham, 147–166. [https://doi.org/10.1007/978-3-319-06520-5\\_10](https://doi.org/10.1007/978-3-319-06520-5_10)
- [72] Diego Riofrio-Luzcando and Jaime Ramirez. 2016. Student Action Prediction for Automatic Tutoring for Procedural Training in 3D Virtual Environments. In *E-Learning, E-Education, and Online Training (Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering)*, Giovanni Vincenti, Alberto Bucchiero, and Carlos Vaz de Carvalho (Eds.). Springer International Publishing, Cham, 200–207. [https://doi.org/10.1007/978-3-319-28883-3\\_25](https://doi.org/10.1007/978-3-319-28883-3_25)
- [73] D. Riofrio-Luzcando, J. Ramirez, and M. Berrocal-Lobo. 2017. Predicting Student Actions in a Procedural Training Environment. *IEEE Transactions on Learning Technologies* 10, 4 (2017), 463–474. <https://doi.org/10.1109/TLT.2017.2658569>
- [74] S. Sahebi, Y. Huang, and P. Brusilovsky. 2014. Predicting Student Performance in Solving Parameterized Exercises. *Lecture Notes in Computer Science* 8474 LNCS (2014), 496–503. [https://doi.org/10.1007/978-3-319-07221-0\\_62](https://doi.org/10.1007/978-3-319-07221-0_62)
- [75] S. Sahebi, Y.-R. Lin, and P. Brusilovsky. 2016. Tensor Factorization for Student Modeling and Performance Prediction in Unstructured Domain. In *Proceedings of the 9th International Conference on Educational Data Mining, EDM 2016*. 502–506.
- [76] John Saint, Dragan Gašević, Wannisa Matcha, Nora Ayu Ahmad Uzir, and Abelardo Pardo. 2020. Combining Analytic Methods to Unlock Sequential and Temporal Patterns of Self-Regulated Learning. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*. ACM, Frankfurt Germany, 402–411. <https://doi.org/10.1145/3375462.3375487>
- [77] Margarete Sandelowski and Julie Barroso. 2007. *Handbook for Synthesizing Qualitative Research*. Springer Publishing Company, New York, NY.
- [78] Carlotta Schatten and Lars Schmidt-Thieme. 2016. Student Progress Modeling with Skills Deficiency Aware Kalman Filters. In *Proceedings of the 8th International Conference on Computer Supported Education*. SCITEPRESS - Science and Technology Publications, Rome, Italy, 31–42. <https://doi.org/10.5220/0005737200310042>
- [79] Robin Schmucker and Tom M Mitchell. 2022. Transferable Student Performance Modeling for Intelligent Tutoring Systems. In *Proceedings of the 30th International Conference on Computers in Education*. Asia-Pacific Society for Computers in Education. The Asia-Pacific Society for Computers in Education, Kuala Lumpur, Malaysia.
- [80] Gayane Sedrakyan, Jonna Malmberg, Katrien Verbert, Sanna Järvelä, and Paul A. Kirschner. 2020. Linking Learning Behavior Analytics and Learning Science Concepts: Designing a Learning Analytics Dashboard for Feedback to Support Learning Regulation. *Computers in Human Behavior* 107 (June 2020), 105512. <https://doi.org/10.1016/j.chb.2018.05.004>
- [81] Andy Smith, Wookhee Min, Bradford W. Mott, and James C. Lester. 2015. Diagrammatic Student Models: Modeling Student Drawing Performance with Deep Learning. In *User Modeling, Adaptation and Personalization*, Francesco Ricci, Kalina Bontcheva, Owen Conlan, and Séamus Lawless (Eds.). Springer International Publishing, Cham, 216–227. [https://doi.org/10.1007/978-3-319-20267-9\\_18](https://doi.org/10.1007/978-3-319-20267-9_18)
- [82] Xiangyu Song, Jianxin Li, Taotao Cai, Shuiqiao Yang, Tingting Yang, and Chengfei Liu. 2022. A Survey on Deep Learning Based Knowledge Tracing. *Knowledge-Based Systems* 258 (Dec. 2022), 110036. <https://doi.org/10.1016/j.knosys.2022.110036>
- [83] J.C.-Y. Sun, S.-J. Yu, and C.-H. Chao. 2019. Effects of Intelligent Feedback on Online Learners' Engagement and Cognitive Load: The Case of Research Ethics Education. *Educational Psychology* 39, 10 (2019), 1293–1310. <https://doi.org/10.1080/01443410.2018.1527291>
- [84] A.A. Supianto and M. Hafis. 2019. GTRAS: Graphical Tracking Activity System for Problem-Posing Learning Process Insights. In *2018 International Conference on Advanced Computer Science and Information Systems, ICACSIS 2018*. IEEE, 231–235. <https://doi.org/10.1109/ICACSIS.2018.8618216>
- [85] Xueying Tang, Zhi Wang, Jingchen Liu, and Zhiliang Ying. 2021. An Exploratory Analysis of the Latent Structure of Process Data via Action Sequence Autoencoders. *Brit. J. Math. Statist. Psych.* 74, 1 (2021), 1–33. <https://doi.org/10.1111/bmps.12203>
- [86] Yun Tang, Zhengfan Li, Guoyi Wang, and Xiangen Hu. 2021. Modeling Learning Behaviors and Predicting Performance in an Intelligent Tutoring System: A Two-Layer Hidden Markov Modeling Approach. *Interactive Learning Environments* 0, 0 (Dec. 2021), 1–13. <https://doi.org/10.1080/10494820.2021.2010100>
- [87] Michelle Taub and Roger Azevedo. 2019. How Does Prior Knowledge Influence Eye Fixations and Sequences of Cognitive and Metacognitive SRL Processes during Learning with an Intelligent Tutoring System? *International Journal of Artificial Intelligence in Education* 29, 1 (March 2019), 1–28. <https://doi.org/10.1007/s40593-018-0165-4>
- [88] Timothy Tuti, Chris Paton, Mike English, and Niall Winters. 2019. Building a Learner Model for a Smartphone-Based Clinical Training Intervention in a Low-Income Context: A Pilot Study. In *Transforming Learning with Meaningful Technologies (Lecture Notes in Computer Science)*, Maren Scheffel, Julien Broisin, Viktoria Pammer-Schindler, Andri Ioannou, and Jan Schneider (Eds.). Springer International Publishing, Cham, 55–68. [https://doi.org/10.1007/978-3-030-29736-7\\_5](https://doi.org/10.1007/978-3-030-29736-7_5)
- [89] Manuel Valle Torre, Marcus Specht, and Catharine Oertel. 2023. Coded Dataset of a Literature Review on Learning Sequence Analysis. <https://doi.org/10.4121/10FC3A10-37AA-4292-818D-468B940B2BB7>
- [90] Manuel Valle Torre, Esther Tan, and Claudia Hauff. 2020. edX Log Data Analysis Made Easy: Introducing ELAT: An Open-Source, Privacy-Aware and Browser-Based edX Log Data Analysis Tool. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge (LAK '20)*. Association for Computing Machinery, New York, NY, USA, 502–511. <https://doi.org/10.1145/3375462.3375510>
- [91] R. Venant, K. Sharma, P. Vidal, P. Dillenbourg, and J. Broisin. 2017. Using Sequential Pattern Mining to Explore Learners' Behaviors and Evaluate Their Correlation with Performance in Inquiry-Based Learning. *Lecture Notes in Computer Science* 10474 LNCS (2017), 286–299. [https://doi.org/10.1007/978-3-319-66610-5\\_21](https://doi.org/10.1007/978-3-319-66610-5_21)
- [92] Jill-Jenn Vie and Hisashi Kashima. 2019. Knowledge Tracing Machines: Factorization Machines for Knowledge Tracing. *Proceedings of the AAAI Conference on Artificial Intelligence* 33, 01 (July 2019), 750–757. <https://doi.org/10.1609/aaai.v33i01.3301750>
- [93] Karen D. Wang, Jade Mai Cock, Tanja Käser, and Engin Bumbacher. 2023. A Systematic Review of Empirical Studies Using Log Data from Open-Ended Learning Environments to Measure Science and Engineering Practices. *British Journal of Educational Technology* 54, 1 (2023), 192–221. <https://doi.org/10.1111/bjjet.13289>
- [94] Yu Wang, Tong Li, Congkai Geng, and Yihan Wang. 2019. Evaluating Student Learning Effect Based on Process Mining. In *Applied Informatics (Communications in Computer and Information Science)*. Hector Florez, Marcelo Leon, Jose Maria Diaz-Nafria, and Simone Belli (Eds.). Springer International Publishing, Cham, 59–72. [https://doi.org/10.1007/978-3-030-32475-9\\_5](https://doi.org/10.1007/978-3-030-32475-9_5)
- [95] Philip H. Winne. 2010. Improving Measurements of Self-Regulated Learning. *Educational Psychologist* 45, 4 (Oct. 2010), 267–276. <https://doi.org/10.1080/00461520.2010.517150>
- [96] Alyssa Friend Wise and David Williamson Shaffer. 2015. Why Theory Matters More than Ever in the Age of Big Data. *Journal of Learning Analytics* 2, 2 (Dec. 2015), 5–13. <https://doi.org/10.18608/jla.2015.22.2>
- [97] Jacqueline Wong, Mohammad Khalil, Martine Baars, Björn B. de Koning, and Fred Paas. 2019. Exploring Sequences of Learner Activities in Relation to Self-Regulated Learning in a Massive Open Online Course. *Computers & Education* 140 (Oct. 2019). <https://doi.org/10.1016/j.compedu.2019.103595>
- [98] Tangjie Wu and Qiang Ling. 2023. Fusing Hybrid Attentive Network with Self-Supervised Dual-Channel Heterogeneous Graph for Knowledge Tracing. *Expert Systems with Applications* 225 (Sept. 2023), 120212. <https://doi.org/10.1016/j.eswa.2023.120212>
- [99] Zhenchang Xia, Nan Dong, Jia Wu, and Chuanguo Ma. 2023. Multi-Variate Knowledge Tracing Based on Graph Neural Network in ASSISTments. *IEEE Transactions on Learning Technologies* (2023), 1–12. <https://doi.org/10.1109/TLT.2023.3301011>
- [100] Wen Xiao, Ping Ji, and Juan Hu. 2022. A Survey on Educational Data Mining Methods Used for Predicting Students' Performance. *Engineering Reports* 4, 5 (2022), e12482. <https://doi.org/10.1002/eng.2.12482>
- [101] X. Xiong, S. Zhao, E.G. Van Inwegen, and J.E. Beck. 2016. Going Deeper with Deep Knowledge Tracing (*Proceedings of the 9th International Conference on Educational Data Mining, EDM 2016*). 545–550.
- [102] Zhenyuan Yang, Shimeng Xu, Changbo Wang, and Gaoqi He. 2022. Skill-Oriented Hierarchical Structure for Deep Knowledge Tracing. In *2022 IEEE 34th International Conference on Tools with Artificial Intelligence (ICTAI)*. 425–432. <https://doi.org/10.1109/ICTAI56018.2022.00069>
- [103] Xiankun Zhang, Yuhu Shang, Yimeng Ren, and Kun Liang. 2023. Dynamic Multi-Objective Sequence-Wise Recommendation Framework via Deep Reinforcement Learning. *Complex & Intelligent Systems* 9, 2 (April 2023), 1891–1911. <https://doi.org/10.1007/s40747-022-00871-x>
- [104] Guojing Zhou, Tetsumichi Umada, and Sidney D'Mello. 2022. What Do Students' Interactions with Online Lecture Videos Reveal about Their Learning?. In *Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '22)*. Association for Computing Machinery, New York, NY, USA, 295–305. <https://doi.org/10.1145/3503252.3531315>