MDPI

Article

Data-driven Analysis of Student Engagement in Time-Limited Computer Laboratories

Luca Cagliero [1,‡]*0000-0002-7185-5247, Lorenzo Canale [1,2†,‡]0000-0002-7556-595X and Laura Farinetti[1,‡]0000-0001-8614-4192

[1] Dipartimento di Automatica e Informatica, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino, Italy; luca.cagliero@polito.it (L.C.); lorenzo.canale@polito.it (L.C.); laura.farinetti@polito.it (L.F.)

[2] Centre for Research and Technological Innovation, Radiotelevisione Italiana (RAI), Via Giovanni Carlo Cavalli 6, 10129, Torino, Italy

* Correspondence: luca.cagliero@polito.it; Tel.: +390110907179 (L.C.)

† Current address: Affiliation 2.

‡ These authors contributed equally to this work.

Abstract: Computer laboratories are learning environments where students learn programming languages by doing practice under teaching assistants' supervision. This paper presents the outcomes of a real case study carried out in our university in the context of a database course, where learning SQL is one of the main topics. The aim of the study is to analyze the level of engagement of the laboratory participants by tracing and correlating the accesses of the students to each laboratory exercise, the successful/failed attempts to solve the exercises, the students' requests for help, and the interventions of teaching assistants. The acquired data are analyzed by means of a sequence pattern mining approach, which automatically discovers recurrent temporal patterns. The mined patterns are mapped to behavioral, cognitive engagement and affective key indicators thus allowing students to be profiled according to their level of engagement in all the identified dimensions. To efficiently extract the desired indicators the mining algorithm enforces ad hoc constraints on the pattern categories of interest. The student profiles and the correlations among different engagement dimensions extracted from the experimental data have shown to be helpful for the planning of future learning experiences.

Keywords: Sequential Pattern Mining; Learning Analytics; Higher Level Education; Engagement

1. Introduction

Laboratories are known to have a primary role in learning activities. Previous research studies (e.g., [1]) have shown that practical activities provide benefits to students in terms of knowledge acquisition, level of engagement, well-being, interaction skills, revision and validation of knowledge competencies. In computer science laboratories often rely on computerized services. They allow students to practice what they have learnt in theory in an interactive way, typically under the supervision of the teaching assistants. Hence, teachers have the opportunity to closely monitor learners in a "natural" learning environment, where they can learn the necessary knowledge by doing. To this purpose, lab assignments typically include exercises of variable complexity thus allowing learning to deal with problems that gradually become similar to the final assessment tasks [2].

Since during computer science laboratories learners commonly work in a controlled environment for a restricted time period, an increasing research interest has been devoted to acquiring, collecting, and analyzing learner-generated data in order measure and monitor students' engagement level during laboratory activities [3]. According to [4], student engagement is *the energy and effort that students employ within their learning community, observable via any number of behavioural, cognitive or affective indicators across a continuum.* Learner engagement can be analyzed under various dimensions, such as (i) the *behavioral*

Citation: Cagliero, L.; Canale, L.; Farinetti, L. Data-driven Analysis of Student Engagement in Time-Limited Computer Laboratories. *Algorithms* **2023**, *1*, 0. https://doi.org/

Received:
Revised:
Accepted:
Published:

aspects, related to observable behavioral characteristics, e.g., the level of effort that students dedicate to learning by participating to the proposed activities and by being involved in the assigned tasks [5], (ii) the *cognitive aspects*, related to students' motivation and investment of thought, mental effort, and willingness to comprehend new ideas and methods [6], and (iii) the *emotional aspects*, related to the affective reactions of the students towards teachers and colleagues [7].

Monitoring and facilitating learning engagement is particularly challenging since it requires to identify the key factors behind students' motivation. Student engagement analytics typically consist of the following steps: first, an appropriate source of information need to be identified. To collect relevant information, previous studies have considered, for instance, data from educational service logs [8], surveys [9], mobile technologies [10], and social networks [11]. Secondly, it entails defining a set of quantitative descriptors of student engagement that are tailored to the specific learning context. Examples of analyzed contexts include, amongst other, MOOCs [9], traditional university-level courses [12], and secondary school lessons [13]. Finally, the acquired data can be analyzed by means of advanced data analytics tools or data mining algorithms in order to extract relevant and promptly usable knowledge. Teachers can exploit the discovered information to facilitate learners' engagement and to improve the quality of the learning activities. Recent surveys on students' engagement and learning technologies [4] acknowledge the need for further research efforts addressing the use of data mining techniques in university-level laboratory activities. The present paper presents a research activities in the aforesaid direction.

This work analyzes the level of engagement of university-level students during computer laboratories on writing database queries in the Structured Query Language (SQL) language. Teaching SQL is widespread in university-level database courses. Computer laboratories are particularly suitable for SQL education because learners could type a the queries solving a list of exercises, progressively submit the draft solutions, and eventually fix them by adopting a trial-and-error approach [14]. We present a case study that we performed in our university, where we set up the laboratory environment and acquired learner-generated data. The designed environment also provides teaching assistants with a prioritized and "democratic" way for giving assistance to students: through an informed environment they can easily spot who is experimenting difficulties according to objective parameters extracted by real-time data collected during the lab. To retrieve data about student engagement, we trace the activities of both students and teaching assistants during the computer lab to analyze the following aspects: (i) the timing and order of access to the given exercises, (ii) the timing of the (potentially multiple) submissions for each assigned exercise, (iii) the submissions' outcome (correct or wrong query), (iv) the requests for teachers' assistants made by the students, and (v) the interventions of the teaching assistants. Therefore, unlike traditional log-based systems, the computer lab scenario allows us to trace key aspects of the learning-by-doing process, such as the sequence of submission successes/failures for a given exercise and the requests for assistance. Acquiring the data described above enables the analysis of a number of key indicators of learner's engagements. To this end, we apply an exploratory sequence pattern mining approach [15] in order to extract temporal patterns from learner-generated data. Patterns describe recurrent and temporally correlated sequences of traced events that can be used to characterize student engagement under multiple perspectives. More specifically, in the present work we will exploit the extracted sequential patterns to answer the following research question: *which kind of information about students' behavioural, cognitive and affective engagement can be extracted from the temporal sequences of the students' activities?* To efficiently extract the desired information, we enforce ad hoc pattern constraints into the sequence mining algorithm. Besides, the collected data have shown to be helpful in addressing issues that are specifically related to the learning experience of the students (e.g., an exercise whose complexity is significantly above average) thus improving the future teaching activities. For example, they help to understand the complexity of the laboratory assignment, evaluate the correctness of the sequence of the proposed exercises, and analyze the impact and effectiveness of the teaching assistance, whenever requested.

The remainder of the paper is organized as follows. Section 2 overviews the related works. Section 3 describes the experimental settings, whereas Section 4 presents the applied methodology. Section 5 reports the analysis of the extracted patterns and discusses the results under the point of view of the students' learning experience. Section 6 focuses on the description of the key engagement indicators extracted by means of sequential pattern mining. It profiles students according to a number of selected behavioral, cognitive and affective engagement dimensions. Furthermore, it analyzes also the correlation among the engagement dimensions extracted from the experimental data. Finally, Section 7 draws the conclusions and future perspectives of this work.

2. Literature review

The use of laboratories in computer science education is established; several studies (e.g., [1,2,16,17]) have highlighted the advantages to have a practical approach to learning, describing facilities, and suggesting best practices. The research community has stressed the importance of cooperation during laboratories. Laboratories are not simply considered as places where a single student interacts with a Personal Computer: their use is primarily concerned with the interaction between students [16,18]. Therefore, studying learners' interactivity inside a lab is particularly useful for improving the effectiveness of learning practices.

The Structured Query Language (SQL) is the most widespread declarative programming language to query relational databases. Due to the overwhelming diffusion of relational DataBase Management Systems, in software engineering and computer science education Structured Query Language (SQL) skills are deemed as fundamental. A systematic review of SQL education is given by [14]. In the early 2000s, most research works related to SQL education were focused on proposing ad hoc tools to support laboratory sessions on SQL query writing (e.g.,[19–21]). Later on, with the growth of Learning Analytics (LA) technologies, the attention of the community has shifted towards the development of smart solutions to acquire, collect, and analyze learner-generated data during SQL laboratories. For example, an established LA challenge is to early predict students' performance [22]. Under this umbrella, the works presented in [23–25] proposed to record students' activities during SQL laboratories in order to obtain inferences related to the upcoming students' performance. More recently, the research community has paid more and more attention to innovative SQL learning paradigms, e.g., blended learning [26,27], game-based learning [28], flipped classrooms activities [29]. The present paper positions itself as a new Learning Analytics study in Higher Education [30], with particular reference to SQL laboratory activities. Unlike [23–25], the focus of the present work is not on predicting students' performance. Conversely, it investigates the use of an exploratory data mining techniques, i.e., sequence pattern mining [31], to characterize and profile learners' activities during SQL laboratory and to describe the cognitive, behavioral, and affective dimensions of student engagement.

In recent years, the parallel issue of fostering student engagement through educational technologies in secondary and higher education has received increasing attention [1,4,8,32]. For example, the authors in [32] analyzed the behavioral engagement of MOOC participants based on both the timing of resource accesses and on the type of explored resources, i.e., video, Self Regulated Learning (SRL) support video, discussion, quiz, assignment, reading. In [8] the authors analyzed click-stream log data related to 89 students of a Freshman English course. They classified students as *surface*, *deep*, or *strategic* according to their engagement level measured in terms of time spent on the Web pages and number of actions made on that pages (detected from reading logs). Some attempts to facilitate students' engagement in secondary education through flipped learning approaches have also been made [4]. An extensive overview of the existing educational technology applications to enhance student engagement in higher education can be found in [1]. Similar to [8,32], in this study we analyze click-stream data in order to monitor students' engagement levels. Unlike [8] we consider a different context of application (i.e., a Higher Education course on databases) and we apply a different methodology for exploring data. Compared to [32], the present work analyzes a different context (i.e., an assisted laboratory activity) and exploits different activity indicators beyond the accesses to a resource, such as the success/fail of a tentative submission of an exercise solution and the interactions with the teaching assistants. The enriched data model enables the study of different learning aspects related to behavioural, cognitive and affective engagement as well. Table 1 enumerates the engagement key indicators that will be addressed in the paper. For each of the selected indicators, the table contains the category (behavioural, cognitive of affective) consistently with the classification proposed in [4], a definition, and a list of related works.

Table 1. Engagement key indicators analyzed in the present paper.

Category	Key Indicator	Definition	References
Behavioural	Persistence	The quality or state of maintaining a course of action or keeping at a task and finishing it despite the obstacles or the effort involved.	[6,33–39]
Cognitive	Concentration	The act of focusing, as, for example, bringing one's thought processes to bear on a central problem or subject.	[6,35,38–40]
	Reflection	A form of theoretical activity directed toward the comprehension of its own acts and the laws by which they are performed. Reflection includes building conclusions, generalisations, analogies, comparisons and evaluations, and also emotional experience, remembering and solving problems. It also includes addressing beliefs for interpretation, analysis, realisation of acts, discussion or evaluation.	[6,35,41,42]
	Understanding	Building complex understanding and meaning rather than focusing on the learning of superficial knowledge.	[6,35,37–41,43]
	Autonomy	A state of independence and self-determination in an individual, a group, or a society.	[6,33,38,39,43,44]
Affective	Confidence	A belief that one is capable of successfully meeting the demands of a task.	[33,35,40,42,44]

3. Experimental setting

To analyze students' activities and engagement in SQL education, the present research work relies on real data collected during educational laboratory sessions. The educational context is a computer lab related to a course on database design and management. The course is offered in the context of a B.S. degree in engineering. All the students are enrolled to the same bachelor degree course, have approximately the same background, and did the practice in the same conditions. The objective of the laboratory activity is to become familiar with the SQL language through a number of proposed SQL exercises, where the student has to write SQL declarative statements to query a relational database.

The computer lab is equipped with 43 workstations, but the course has about 650 enrolled students; for this reason, students were divided in groups and participate to a 90-minute lab session. The task consisted in solving 13 proposed exercises through with educational tool that supports them and records all the related events. The first 4 exercises just require the knowledge of the basic SQL syntax `SELECT ... FROM ... WHERE ... ORDER BY`, the subsequent 4 exercises require a more advanced understanding of SQL grouping operators (`GROUP BY ... HAVING ...`), whereas the remaining ones mainly focus on nesting SQL queries using Table Functions and the `IN, EXISTS, NOT IN, NOT EXISTS` operators.

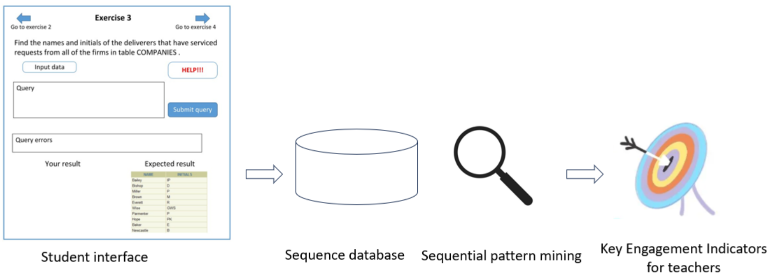
The students' user interface proposes one exercise at a time, with the problem statement, the associated relational database schema and the table representing the expected correct results. The students enter their tentative query and the Oracle DBMS [45] executes it, providing the feedback that is shown to the learners. Besides the DBMS messages (useful for understanding query errors), when the query is syntactically correct the environment compares the executed result with the expected result, thus highlighting possible semantic errors.

Through the user interface students can also ask for the teaching assistant's intervention; the environment records both help requests and interventions.

Participation to labs was optional (even tough highly encouraged). Therefore not every student participated to the lab experiment. For this study we collected the data about 215 students, considering only those who accessed at least one exercise.

4. Materials and Methods

Figure 1. Designed pipeline.



The pipeline of analysis designed for studying student engagement in SQL education during computer laboratory consists of three main steps (see Figure 1). Firstly, the data are acquired through computer laboratory interface. Then, data are tailored to an appropriate **sequence database**, which incorporates all the necessary information. Secondly, a subset of relevant temporal patterns is extracted using an established **sequential pattern mining** approach [15]. Pattern extraction aims at automatically extracting recurrent subsequences of temporally correlated events related to student engagement. Lastly, a set of **Key Engagement Indicators (KEIs)** (see Table 1) are computed on top of the extracted patterns. KEI exploration can help teachers to monitor and facilitate learner engagement from multiple perspectives.

In the following sections, the above-mentioned steps will be thoroughly described.

4.1. Preliminaries

We first define the preliminary concepts of sequence and sequential database. sequential pattern mining in compliance with [46].

Let I be a set of all items. An itemset is a subset of items in I. A sequence s is an ordered list of itemsets. A sequence s is denoted by $\langle s_1 s_2 \dots s_l \rangle$ where $s_j \in I, 1 \leq j \leq l$, is an itemset. s_j is also denoted by *element* of the sequence, consisting of a set of items $(x_1 x_2 \cdot x_m)$, where $x_k \in I, 1 \leq k \leq m$. For the sake of brevity, hereafter we will omit the brackets when m=1. An item occurs at most once in an element of a sequence, but can occur multiple times in different elements of the same sequence. A l-sequence, i.e., a sequence of length l is a sequence whose as the number of instance of occurring items is l. α= $\langle \alpha_1 \alpha_2 \dots \alpha_l \rangle$ is a subsequence of another sequence β= $\langle \beta_1 \beta_2 \dots \beta_l \rangle$, denoted by $\alpha \sqsubseteq \beta$, if there exist integers $1 \leq j_1 \leq j_2 \dots \leq m$ such that $\alpha_1 \subseteq \beta_{j_1}, \alpha_2 \subseteq \beta_{j_2}, \dots, \alpha_n \subseteq \beta_{j_n}$.

A sequence database S is a set of tuples $\langle sid, s \rangle$, where sid is the sequence identifier and s is a sequence. A tuple $\langle sid, s \rangle$ contains a subsequence α if $\alpha \sqsubseteq s$. The absolute support of subsequence α in S, denoted by $\sup_S(s)$ is the number of tuples containing α. The relative support is the fraction of tuples containing α.

4.1.1. Sequential pattern mining

Given a sequence database S and a minimum support threshold `minsup`, the sequential pattern mining task entails extracting all the subsequences α in S whose $\sup_S(\alpha) \geq$ `minsup`, i.e., it focuses on discovering all the *frequent subsequences* in the sequence database.

Whether the occurrences of the sequence elements are timestamped, i.e., t_j is the timestamp at sequence s_j occur, we can enforce additional constraints into the sequential pattern mining process (beyond enforcing support threshold):

- `mingap`: `mingap` the minimum time gap between consecutive elements of a sequence.
- `maxgap`: it indicates the the maximum time gap between consecutive elements of a sequence.
- `maxwinsize`: the maximum temporal duration of the overall sequence.

When not otherwise specified, time gaps and window sizes are expressed in minutes.

By varying the values of `mingap`, `maxgap`, and `maxwinsize` it is possible to focus the exploration on sequences with varying temporal periodicity.

4.2. Data model

We introduce the notation used throughout the section below.

- **Participating students** ($\mathcal{S}$): set of students who participated to a SQL laboratory (i.e., in our experiments, 215 students).
- **Lab duration** (D): The time span corresponding to lab development (i.e., a 90-minute time window, in our experiments).
- **Time window** (TW): A time span at a finer granularity than D (e.g., a 5-minute time span).
- **Events** ($\mathcal{E}$): The set of events of interest occurred in the SQL laboratory. An event $e \in \mathcal{E}$ that occurred at an arbitrary time point $t_e \in D$ and involved a specific student $s \in \mathcal{S}$.

The analysis focuses on the most relevant temporal correlations between the events that occurred in the labs and are relative to the same student. Each event describes either a specific action made by the student (e.g., access to a new exercise), an achievement (e.g., exercise solved), a request for assistance, or an assistance intervention. As discussed later on, the selected events are deemed as relevant to quantify the key engagement indicators under analysis. For our convenience, hereafter each event will be represented by a symbol consisting in the number of the exercise surrounded by a colored shape that describes the type of the event. Specifically,

- the symbol 1 represents an access to exercise 1;
- the symbol $\boxed{1}$ represents the submission of a correct solution for exercise 1;
- the symbol 1 represents the failure of exercise 1;
- the symbol 1 represents an assistance request for exercise 1;
- the symbol 1 represents assistance for exercise 1.

Since the main goal of the study is to quantify the engagement key indicators of the students attending a SQL laboratory using the most representative temporal sequences of events, we rely on a event data model consisting of a sequence database [31] and described in Section 4.1. Specifically, each symbol describing an event is an item and each subsequence is an ordered list of single events (or event sets) associated with a given student.

For example, the subsequence $\langle 1\ 1\ \boxed{1} \rangle$ represents a student that accesses exercise 1, fails it, and then subbits the correct solution.

4.3. The CSpade Algorithm

The CSpade algorithm [47], whose pseudocode is given in Algorithm 1, extracts all subsequences satisfying the input constraints by adopting a prefix-based strategy. The key idea is to decompose the original problem into smaller sub-problems using equivalence classes on frequent sequences. Each equivalence class be solved independently and likely fits in main-memory. The enumeration step is known to be the most computationally intensive one and is traditionally performed via Breadth-First Search (BFS) or Depth-First Search (DFS) [47]. However, as discussed later on in Section 4.4, we envisage a further algorithmic optimization.

Algorithm 1 CSpade [47]

Require: DB, `minsup`, `mingap`, `maxgap`, `maxwinsize`
Ensure: `Sequences` SQ
  $F_1 \leftarrow x$ {Frequent elements}
  $F_k \leftarrow$ {Frequent sequence of k elements}
  for k=2; $F_k \neq$; k=$k+1$ **do**
    Enumerate all the frequent sequences via BFS/DFS ▷ This step will be further optimized (see Section 4.4)
    $C_k \leftarrow$ {Candidate sequences of length k}
    while $s \in DB$ **do**
      for $c \in C_k$ **do**
        Update c.support, c.size, c.gap
      end for
    end while
    $F_k \leftarrow \{c \in C_k | c$ satisfies all input constraints$\}$
  end for

4.4. Computation and Analysis of Engagement Key Indicators

Teachers explore the sequential patterns extracted at the previous step to gain insights into students' engagement in the SQL computer laboratories.

The student-related events considered in this study (see Section 4.2) are exploited to analyze student involvement, motivation, and willingness to comprehend the fundamentals behind the SQL language. Specifically, the aim is to analyze the sequence database in order to characterize the behavioral, cognitive and affective engagement levels of the students who participated to the laboratories.

The occurrence of single events (e.g., the access to a specific exercise) is not relevant enough to profile students according to their engagement level because it is likely to be related to the occurrence of other events occurred in the past, potentially regarding different event types and exercises. Hence, the present analysis relies on the extraction of sequential patterns, which represent the most significant temporal correlation between the occurrences of multiple events. The idea behind is to capture the most interesting temporal relationships between correlated events and get actionable knowledge about student activities, involvement, and motivation.

Based on the characteristics of the contained events, the extracted sequential patterns can be classified as follows.

- **Access patterns**: This type of patterns comprises all the sequences whose elements are *exclusively* composed of events of type *access to exercise*. Since students (i) are provided with a ordered list of exercises, (ii) have no time limits to solve an exercise, (iii) can move back-and-forth in the exercise list according to their preferences, exploring access patterns allows teachers to understand the way students deal with the laboratory exercises as well as to analyze the time spent on each exercise.
- **Successful patterns**: This pattern category includes all the sequences whose elements comprise both access and successful attempts for the same exercise. They are deemed as relevant to explore both the level of complexity of the provided exercises and the level of competence of the students.
- **Assistance request patterns**: This type of patterns includes all the sequences that comprehend a request for assistance.

- **Assistance intervention patterns**: This type of patterns consists of all the sequences that comprehend an intervention of the teaching assistant. Together with the assistance request patterns, they provide interesting insights into the ability of the students to work in autonomy. They allow us also to identify the most common situations when students ask for help, and to study the impact of the intervention of a teacher assistance on the development of the current and following exercises.
- **Error patterns**: This pattern type comprises all the sequences whose elements include events of type *Wrong submitted query for a given exercise*. They can be exploited to identify the exercises generating major difficulties, to cluster students based on their level of competence, as well as to monitor the progresses of the students across the practice (e.g., to understand whether the trial-and-error approach actually works or not).
- **Time-constrained patterns**: This class of patterns consists of all the sequences extracted by enforcing either a minimum/maximum gap between each element of the sequence or a maximum sequence duration (i.e., the elapsed time for the occurrence of the first element and those of the last one). Unlike all the previous pattern types, they give more insights into the timing of specific event. They can be exploited to analyze the timing of the activities and the responsiveness of a student (e.g., the time needed to submit the first query, the time needed to resubmit a query after a failure, the overall time spent in solving an exercise).

As detailed in Table 2, the above-mentioned pattern categories are mapped to the engagement key indicators reported in Table 1.

Algorithmic optimization based on KEI information

To efficiently extract the Key Engagement Indicators we enforce further mining constraints deeply into the candidate sequence generation process (see Algorithm 1). Specifically, similar to [48] we use regular expressions to early discard ordered sequences of elements that do not meet any of the categories reported in Table 2. This prevents the generation and evaluation of an unnecessary large set of candidate sequences, among which many of them are potentially not relevant to students' engagement level analysis.

5. Results

Multiple sequential pattern mining sessions were run on the sequence database acquired during the SQL laboratories of a B.S. course hold in our university (see Section 4.2). The mined sequential patterns are explored in order to evaluate the effectiveness of the proposed methodology in supporting and monitoring students' engagement levels.

The experiments were run on a machine equipped with Intel(R) Core(TM) i7-8550U CPU with 16 GB of RAM running Ubuntu 18.04 server. To extract sequential sequential patterns, we used the CSpade algorithm implementation provided by the respective authors. Multiple mining sessions were run, by varying the `minsup` value to extract sequential patterns without time constraint, and by varying `minsup`,`mingap`, `maxgap`, and `maxlen` to mine time-constrained patterns.

5.1. Access patterns

These patterns describe the timing of the students' accesses to the proposed exercises during the SQL laboratory session. A sample of the extracted sequences is reported in Table 3, with the relative support value (percentage of students that satisfy the specific sequence). Based on the sequences belonging to this pattern type, students can be clustered in two groups based on their profile of accesses to the proposed exercises:

- *Students using sequential patterns*: this cluster consists of the students who accessed the exercises in the proposed sequence (from exercise 1 to 13).
- *Students using out-of-order patterns*: it groups the students who follow a non-sequential order in accessing the assigned exercises.

Table 2. Key Engagement Indicators and associated patterns.

Key engagement indicator	**Pattern type**	**Comments**
Persistence	Access patterns	These patterns indicate the persistence of the student on a specific SQL exercise. They discriminate between students adopting a sequential approach, i.e., they address the exercises according to the given order, and those adopting an out-of-order approach, i.e., they reconsider the previously accessed exercises by going back-and-forth between the provided exercises.
Concentration	Successful patterns	These patterns indicate the tendency of a student to focus on a specific exercise until a solution has been found. They discriminate between the students adopting a try-until-successful approach, which entails insisting on the same exercise until a solution has been found, and those adopting a move-to-the-next-exercise approach, which entails jumping to other exercises before solving the current one.
Confidence	Assistance requests patterns	These patterns indicate the level of self-confidence of the students in solving the proposed exercises on their own. They allow to discriminate between the students who are used to ask for help during the computer lab session and the students who generally try to solve the main issues on their own.
Reflection	Errors patterns	These patterns indicate the ability of the student to learn from her/his mistakes. They discriminate between students with a strong reflective attitude, who carefully analyze each error in order to minimize the error rate at the next submission, and students with less reflective attitude, who adopt a trial-and-error approach thus spending a very limited amount of time in understanding the reasons behind the errors before the next submission.
Understanding	Time-constrained patterns	These patterns highlight the timing of a student working on an exercise. They provide useful hints to answer to questions such as: (i) What is the (overall) average time spent by students on each exercise? (ii) What is the average time spent on an exercise prior to submit a query? (iii) What is the average time needed to resubmit a new solution after a failure? (iv) What is the average time needed to solve an exercise?
Autonomy	Assistance intervention patterns + successful patterns	Assistance intervention patterns highlight the effect of an intervention by a teaching assistant on the solution of the current exercise and of the following ones. Together with the successful patterns, they discriminate students that are able to solve exercises autonomously and those who need extra explanations.

Sequential patterns reveal that most of students consecutively accessed the first 5 exercises. However, as the exercise number increases the pattern support decreases. For example, it decreases by 4% from `A1` to `A2` and by 9% (179 students) from `A2` to `A3`. Besides, the frequency count halves from `A3` to `A4`. This result reflects the actual complexity of the proposed exercises: teaching assistants confirmed that the perceived complexity of exercise 5 was higher than expected. It should be noted that the application used by the students during the laboratory allowed them to access a specific exercise only after all previous ones are accessed. This is the reason why skipped exercises never occurred in these patterns[1].

Out-of-order patterns reveal the students who came back to a previous exercise. In [49] the authors highlighted the usefulness of “design by copying” practice, whereas in [50] the authors paid attention to the "we do as we did in the previous exercise" thinking in learning practice. These behaviours occur also in this learning context and explain why the students are used to come back to the previous exercises; most of students face the SQL language practice for the first time and they are not yet familiar with the subject.

[1] An exercise is considered as *skipped* when the student did not access it.

Table 3. Access patterns.

Id	Pattern	Students	Students (%)
Sequential patterns			
A1	1 2 3	204	94.9
A2	1 2 3 4	196	91.2
A3	1 2 3 4 5	179	83.3
A4	1 2 3 4 5 6	98	45.6
Out-of-order patterns			
A5	1 2 1	43	20.0
A6	2 3 2	39	18.1
A7	3 4 3	42	19.5
A8	4 5 4	42	19.5
A9	5 6 5	36	16.7

Table 3 shows that *out-of-order* patterns are almost equally spread over the first 6 exercises; in fact the support value does not show any significant variation, as happened for the *sequential* sequences. Conversely, it slightly varies between 16.7% (36 students) and 20% (43 students).

The differences between sequential and out-of order sequences are likely to be related to the *"Persistence"* indicator of behavioural engagement. This aspect will be discussed later on (see Section 6).

5.2. Successful patterns

This pattern type describes the sequences that contain accesses and successful query submissions. The top ranked sequences (in order of decreasing support value) are reported in Table 4.

We can differentiate between *sequential patterns* and *out-of-order patterns* even in this case; the first ones reveal the students that accessed an exercises only after having solved all the previous ones. 81.4% of the students who solved the first 2 exercises sequentially (pattern `S1`, $\mathtt{sup}_{\mathtt{perc}}\mathtt{(S1)} = 81.4\%$) did the same also for exercise 3 (pattern `S2`, $\mathtt{sup}_{\mathtt{perc}}\mathtt{(S2)} = 68.4\%$). Skipping exercise 3 is therefore a relatively rare condition. On the contrary, only 61.9 % of the students that completed the third exercise succeeded also in the fourth one (pattern `S3`, $\mathtt{sup}_{\mathtt{perc}}\mathtt{(S3)} = 42.3\%$). The `sup(S4)` (93 students) is almost equal to `sup(S3)` (91 students): only 1% (2 students) who solved the first four exercises did not solve exercise 1.

By comparing `S2` with the access pattern `A1`, it appears that 27% of the students (58) who accessed the first three exercises did not solve at least one of them or even many of them; such a percentage increases (46.4%, 105 students) while considering also the fourth exercise (hence comparing `S3` with `A2`). This means than more than half of the students who accessed the first four exercises failed at least one of them.

The *out-of-order patterns* do not show the students who accessed an exercises without solving the previous ones, as one might think: they only show the students that accessed and solved the exercises contained in the pattern, without explicitly revealing that they did not solve the exercises that do not appear in the pattern. This is mainly due to the peculiar characteristics of the sequential patterns [15]. This means that all sequences that contain `S2` also contain `S6`, and therefore we can derive the percentage of students who solved exercises 1 and 3, but not exercise 2, by computing $\mathtt{sup(S6)} - \mathtt{sup(S2)} = 4.6\%$.

Table 4. Successful patterns.

Id	Pattern	Students	Students (%)
	Sequential patterns		
S1	① $\boxed{1}$ ② $\boxed{2}$	175	81.4
S2	① $\boxed{1}$ ② $\boxed{2}$ ③ $\boxed{3}$	147	68.4
S3	① $\boxed{1}$ ② $\boxed{2}$ ③ $\boxed{3}$ ④ $\boxed{4}$	91	42.3
S4	② $\boxed{2}$ ③ $\boxed{3}$ ④ $\boxed{4}$	93	43.3
S5	③ $\boxed{3}$ ④ $\boxed{4}$	94	43.7
	Out-of-order patterns		
S6	① $\boxed{1}$ ③ $\boxed{3}$	157	73.0
S7	① $\boxed{1}$ ② $\boxed{2}$ ④ $\boxed{4}$	102	47.4
S8	① $\boxed{1}$ ③ $\boxed{3}$ ④ $\boxed{4}$	92	42.8
S9	② $\boxed{2}$ ④ $\boxed{4}$	104	48.4

In a similar way, we can compute `sup(S7)` $-$ `sup(S3)` $= 5.1\%$, `sup(S9)` $-$ `sup(S4)` $= 5.1\%$, and `sup(S8)` $-$ `sup(S3)` $= 0.5\%$. The latter result clearly indicates that the difference between the students who solved exercises 1, 3 and 4 and the ones who solved all the four exercises is only 1 student. Therefore, the second task was the easiest one for the students who solved these subset of exercises.

The successful pattern sequences can be related to the *"Concentration"* key indicator of cognitive engagement, as discussed later on in Section 6.

5.3. Assistance patterns

This pattern category helps to analyze the students' requests for help and the assistants' responses. The patterns are divided into 2 subcategories: *Assistance request patterns* and *Assistance intervention patterns*. The former one reveals when and how often students ask for help, whereas the latter discloses when and how often assistants take action and quantifies the consequent effect. Table 5 reports the top ranked patterns separately for each subcategory.

Pattern `H1` shows that some students asked for help more than once. This particular situation happened only for exercise 1: the students' attitude in case of the first exercise is different with respect to the next exercises, considering also that most of students requested assistance just once in the whole lab session.

86% of the students who requested assistance (80 students out of 93) then solved it (pattern `H2`); by comparing `H2` and `H4` it turns out that 61 of them solved it after the assistance, whereas 19 of them succeeded autonomously.

The difference between students who succeeded after requesting assistance (`H3`, `sup(H3)`=54) and the students who succeeded after assistant interventions (pattern `H5`,`sup(H5)`=52) is less significant for exercise 2: only 2 students who asked for help solved the exercise autonomously. Notably, in exercise 3 all students that succeeded after requesting help have been assisted.

Patterns `H10`, `H11` and `H12` show the number of errors after assistants' interventions for exercises 1, 2 and 3 respectively. As the exercise number increases, the support decreases; this is because exercise 2 and 3 generally were perceived as easier than exercise 1 (this situation will clearly emerge later on in the analysis of the time constrained patterns). Note also as the exercise identifier increases the number of students who accessed it decreases (as previously discussed in the *Accesses patterns* analysis).

Table 5. Assistance patterns.

Id	**Pattern**	**Students**	**Students (%)**
Assistance request patterns			
H1	1 1 1	35	16.3
H2	1 1 $\boxed{1}$	80	37.2
H3	2 2 $\boxed{2}$	54	25.1
H4	3 3 $\boxed{3}$	36	16.7
Assistance intervention patterns			
H5	1 1 $\boxed{1}$	61	28.4
H6	2 2 $\boxed{2}$	52	24.2
H7	3 3 $\boxed{3}$	36	16.7
H8	1 1 1	52	24.2
H9	2 2 2	49	22.8
H10	3 3 3	32	14.9
H11	1 1 1 $\boxed{1}$	48	22.3

Pattern `H10` identifies the students who received assistance, committed errors and finally succeeded in exercise 1; by comparing the support value of such a pattern with those of `H4`, we can conclude that only 13 students succeeded immediately after receiving help.

The pattern of type "intervention-error-success" occurs only for exercise 1. For the next exercises the minimum support threshold was not reached. Both *Request effectiveness* and *Assistance effectiveness* decrease as the exercises identifiers increase because the exercises become more difficult and the effects of assistants' interventions are probably less evident in the very short-term.

The assistance patterns can be related to the *"Confidence"* key indicator of cognitive engagement (assistance request patterns) and to the *"Autonomy"* key indicator of affective engagement (assistance intervention patterns), as analyzed later on in Section 6.

5.4. Error patterns

This type of patterns is useful for describing the way students react to errors. We distinguish between *single errors* patterns, which give a general overview about error distribution, and *repeated errors* patterns, which describe how many time an error occurred. The most frequent sequences of both categories are reported in Table 6.

The support value of the *single errors* patterns from `E1` to `E6` show the number of students who solved a specific exercise after making at least one error. The *Students (%)* column in the table shows that most of the students who initially failed, succeeded in the first three exercises; on the contrary, this is not true for exercises 4 and 5. *Students (%)* tends to decrease as the exercise number increase, because the queries become gradually more and more complex.

Pattern `E7` indicates that 59.5% of students made at least one mistake for the exercises from 1 to 3. Many errors are relative to these exercises, considering that 94.9% accessed them (see pattern `A1`). Pattern `E8` reveals a similar behaviour: in fact, the percentage of students who committed errors in all the first four exercises is high (47.9%).

Patterns `E9`, `E10` and `E11` show that at least half of the students committed errors before succeeding in at least one of the first three exercises, and this is coherent with the fact that students are currently learning the SQL language. In [51] the authors stated that most of query errors are simply trial and error inputs as incomplete attempt derived by lack of attention and syntax understanding. Trial and error schema is quite common method in SQL learning.

The *repeated errors* patterns confirm this behaviour; in fact, patterns from E12 to E21 highlight that many wrong queries are relative to the same exercise, whereas patterns E22 and E23 show that this may happen more than once for the same student.

The difference between single errors and repeated errors patterns can be related to the *"Reflection"* key indicator of cognitive engagement, as discuses later on in Section 6.

Table 6. Error Patterns

Id	Pattern	Students	Students (%)
	Single errors		
E1	1 1 $\boxed{1}$	169	78.6
E2	2 2 $\boxed{2}$	141	65.6
E3	3 3 $\boxed{3}$	128	59.5
E4	4 4 $\boxed{4}$	75	34.9
E5	5 5 $\boxed{5}$	35	16.3
E6	6 6 $\boxed{6}$	36	16.7
E7	1 1 2 2 3 3	128	59.5
E8	1 1 2 2 3 3 4 4	103	47.9
E9	1 1 $\boxed{1}$ 2 $\boxed{2}$ 3 $\boxed{3}$	128	59.5
E10	1 $\boxed{1}$ 2 2 $\boxed{2}$ 3 $\boxed{3}$	116	54.0
E11	1 $\boxed{1}$ 2 $\boxed{2}$ 3 3 $\boxed{3}$	117	54.4
	Repeated errors		
E12	1 1 1 1	135	62.8
E13	2 2 2 2	118	54.9
E14	3 3 3 3	102	47.4
E15	4 4 4 4	81	37.8
E16	5 5 5 5	90	41.9
E17	1 1 1 1 1 1 1	70	32.6
E18	2 2 2 2 2 2 2	82	38.1
E19	3 3 3 3 3 3 3	49	22.8
E20	4 4 4 4 4 4 4	41	19.1
E21	5 5 5 5 5 5 5	55	25.6
E22	1 1 1 2 2 2	110	51.2
E23	1 1 1 1 2 2 2 2 2 2	72	33.5

5.5. Time constrained patterns

Time constrained patterns are exploited to answer specific questions related to the timing of the laboratory activities. They can be related to the *"Understanding"* indicator of cognitive engagement, as discussed later on in Section 6.

We set `mingap` to 10 and varied the `maxgap` values from 10 seconds to 5 minutes (i.e., 10s, 60s, 120s, 180s, 240s, 300s). Hence, here we focus on small time intervals to capture short-term student behaviors. The extracted patterns are reported in Table 7.

Table 7. Time constrained patterns.

Id	Pattern	Students	Students (%)
	`maxgap`=60		
`T1`	① ◇1	30	14.0
	`maxgap`=120		
`T2`	① ◇1	47	21.9
`T3`	② ◇2	35	16.3
`T4`	③ ◇3	27	12.6
	`maxgap`=180		
`T5`	② $\boxed{2}$	28	13.0
`T6`	① ◇1	59	27.4
`T7`	② ◇2	56	26.0
`T8`	③ ◇3	38	17.7
`T9`	④ ◇4	32	14.9
`T10`	⑤ ◇5	24	11.2
`T11`	⑥ ◇6	30	14.0
`T12`	⑦ ◇7	22	10.2
	`maxgap`=240		
`T13`	② $\boxed{2}$	43	20.0
`T14`	③ $\boxed{3}$	30	14.0
`T15`	④ $\boxed{4}$	31	14.4
	`maxgap`=300		
`T16`	② $\boxed{2}$	46	21.4
`T17`	③ $\boxed{3}$	30	14.0
`T18`	④ $\boxed{4}$	38	17.7
`T19`	⑥ $\boxed{6}$	23	10.7
`T20`	① ◇1 $\boxed{1}$	47	21.9

Most of the attempts submitted in the very first minutes failed. 30 students who accessed exercise 1 made a mistake in less than one minute (see pattern `T1`) . By increasing the maximum gap threshold to 2 minutes the number of failures for exercise 1 increases and some wrong queries for exercises 2 and 3 start to appear (patterns `T4` and `T3`). By setting `maxgap` to 180, access-error patterns appear for most exercises (from `T6` to `T12`), revealing that the practice to try to submit a solution very quickly is quite popular; in addition, `T5` shows that 13% of students solved exercise 2 in less than 3 minutes (this particular exercise is the one that was solved, on average, in the shortest amount of time). Even though the required competences are slightly more advanced than in the previous exercise, students have already become familiar with the learning environment.

By increasing the maximum gap threshold to 4 minutes, the access-success patterns related to exercise 2 become more frequent (pattern `T13`), and similar patterns occur for exercise 3 and 4 (pattern `T14` and `T15`). When the maximum threshold is set to 5 minutes the same pattern occurs for exercise 6 too (see pattern `T19`). Access-success patterns for exercises 1 and 5 do not appear when `maxgap` is set to 300, since they required more than 5 minutes to be solved.

Patterns `T16` to `T19` show the percentage of students who solved exercises 2, 3, 4 and 6 in less than five minutes: considering such a time constrain, exercise 2 was solved by 21.4%, exercise 3 by 14%, exercise 4 by 17.7% and exercise 6 by 10.%.

By setting *mingap* to 600 and *maxgap* to 900 (time intervals between 10 and 15 minutes) the extracted patterns (reported in Table 8) are all related to the exercises 1, 3 and 5. This shows that these are the exercises for which the students encountered most of the issues.

The difficulty level experienced by the students is not always directly related to the actual difficulty level of the exercises, because other factors can influence, such as the familiarity with the learning environment that plays an important role when the approach is mainly a trial-and-error one.

Table 8. Patterns for interval 10-15 minutes.

Id	Pattern	Students	Students (%)
`T21`	1 1	97	45.1
`T22`	3 3	81	37.7
`T23`	5 5	74	34.4

To detect the lab activities that required longer time, here we set `mingap` to 1800 (30 minutes) and we did not enforce any `maxgap` constraint. Table 9 reports the extracted patterns.

Table 9. Long time patterns.

Id	Pattern	Students	Students (%)
`L1`	1 2	34	15.8
`L2`	1 5	158	73.4
`L3`	1 4 6	24	11.2
`L4`	3 5	82	38.1

15.8% of students spent more than 30 minutes on exercise 1 before accessing exercise 2 (pattern `L1`). This points out once again the problems discussed previously about exercise 1. Another interesting pattern is `L2`: it reveals that 73.5% of students spent at least 30 minutes before accessing exercise 5 after having accessed exercise 1. Considering that the laboratory session lasted 90 minutes, consisted of 13 exercises of increasing difficulty, students proceeded very slowly (notice however that they are not supposed to finish all exercises in the lab, but to finish them as homework). The comparison between `L2` and pattern `A2` shows that only 21 students accessed exercise 5 after 30 minutes (9.7% of all students, 10.3% of those who accessed exercise 1).

Pattern `L3` confirms the difficulties in solving the first exercises of the lab: 24 students (11.2%) who accessed exercise 1 accessed exercise 4 after at least half an hour and exercise 6 after another 30 minutes. 82 students (38.1%) who accessed exercise 3, accessed exercise 5 after 30 minutes (pattern `L4`); this means that solving both exercise 3 and 4 took a long time. Considering the difficulty rank deduced before, and the error patterns in Table 6, this is mainly due to the high number of errors and the time spent on exercise 3.

5.6. Discussion

The extracted patterns can be used to gain insights into the students' learning experience during the SQL laboratory sessions. Very few students completed all the assigned exercises: most of them completed only the first six exercises. The results confirm that the proposed practice was way too long for a 90-minute session. Teachers' objective, in fact, were to challenge the students with more exercises than those strictly requested in order to encourage them to complete the practice at home.

Access patterns show that as exercise number increases the number of students accessing to it decreases, because most of them are struggling on the previous ones, whereas *Successful patterns* and *Error patterns* show that few students who solved exercise 4 passes all first four exercises; these findings reveal the general difficulty in solving the first part of the lab session.

In a time interval of 5 minutes after the access to the exercise (see Table 7) a significant number of students could solve only exercises 2, 4, 3 and 6. Exercises 1, 3 and 5 were the ones where students had more problems (see Table 8). Besides, Table 9 shows that about 16% of students spent more than 30 minutes on exercise 1 before accessing exercise 2, and that about 3 out of 4 students spent at least 30 minutes before accessing exercise 5 after having accessed exercise 1. A difficulty disparity between exercises 2, 3 and 4 and exercises 1 and 5 is therefore evident. About exercise 1, this is understandable because most students used the learning environment for the first time, and this was also the first time they practiced SQL. Exercise 5 caused many problems for most of the students because it introduced new SQL language structures.

Assistance patterns show that the requests for help and the assistants interventions are usually useful for solving the exercises, and that the students succeeded in most cases after being helped. Students were used to ask for help after a few minutes from the exercise accesses, and often many students asked for assistance simultaneously; this caused a waiting time up to 10 minutes before being assisted. In addition, they rarely required assistance twice for the same exercise. The assistants usually intervened after 10 minutes, due to the high number of assistance requests. In addition to the startup delay, there are some specific exercises (especially the number 5) that required long time to be solved. Some of the students solved the exercise before the assistant interventions (especially for exercise 1).

In general students submitted several wrong queries before the correct one, showing a trial-and-error approach that is typical for the laboratory session in computer science courses.

Through sequential pattern analysis, teachers could reinforce the lab experience by considering the discovered issues. First of all, an introduction of the lab environment could be suitable for limiting the startup problems; some exercises could be solved step-by-step by the assistants to prepare the students to the autonomous work. The sequence of the proposed exercises could also be modified to better reflect the students' perceived difficulties.

6. Engagement analysis

The extracted sequences can be conveniently used to describe the engagement characteristics of the students who participated to the SQL laboratory sessions. Specifically, we consider the Key Engagement Indicators described in Table 1 and the association between KEIs and sequential pattern types reported in Table 2 (see section 4.4). In the following, we present both the results of the students' profiling step according to their engagement characteristics and the outcomes of the correlation analysis between different KEIs.

6.1. Students' profiling

Students can be described according to their level associated with each of the six KEIs. For the indicators *Concentration*, *Reflection* and *Autonomy* we define two levels (*High* or *Low*), whereas for the *Persistence*, *Confidence* and *Understanding* we exploit a three-level categorization (*High*, *Medium* or *Low*). Table 10 contains the details of the sequences used to assign the students to a specific level of a given KEI. 561 562

Table 10. Sequences used to assign the students to a specific level of a given key engagement indicator.

Key engagement indicator	**Pattern type**	**Patterns**	**Indicator level**	**Comments**
Persistence	Access patterns	Student satisfies at least one sequence in set (A1-A4) but no sequence in set (A5-A9)	*High persistence*	Only sequential access patterns
		Student satisfies at least one sequence in set (A1-A4) and at least one in set (A5-A9)	*Medium persistence*	Mixed access patterns
		Student satisfies at least one sequence in set (A5-A9) but no sequence in set (A1-A4)	*Low persistence*	Only out-our-order access patterns
Concentration	Successful patterns	Student satisfies at least one sequence in set (S1-S9)	*High concentration*	Stays focused on an exercise until it is solved correctly
		Student does not satisfy any sequence in set (S1-S9)	*Low concentration*	Does not stay focuses on an exercise until it is solved correctly
Confidence	Assistance requests patterns	Student does not satisfy any sequence in set (H1-H4)	*High confidence*	No request for help
		Student satisfies at least one sequence in set (H2-H4) but not sequence H1	*Medium confidence*	Maximum one request for help per exercise
		Student satisfies sequence H1	*Low confidence*	Multiple requests for help for the same exercise
Reflection	Errors patterns	Student satisfies at least one sequence in set (E1-E6) or in set (E9-E11) but no sequence in set (E12-E23)	*High reflection*	Single error before the correct solution
		Student satisfies at least one sequence in set (E12-E23)	*Low reflection*	Repeated errors
Understanding	Time-constrained patterns	Student satisfies sequence T5 or at least one sequence in set (T13-T15) but no sequence in set (T16-T20)	*High understanding*	Correct solution in a short amount of time (e.g. 2-3 minutes)
		Student satisfies at least one sequence in set (T16-T20)	*Medium understanding*	Correct solution in an higher amount of time (e.g. <5 minutes)
		Student satisfies at least one sequence in set (T2-T4) or in set (T6-T12) but not sequence T5 and no sequence in set (T13-T15) or in set (SI1-SI5)	*Low understanding*	No correct solution in a given amount of time (e.g. 5 minutes)
Autonomy	Assistance interventions patterns + successful patterns	Student satisfies at least one sequence in set (S1-S9) but no sequence in set (H5-H11)	*High autonomy*	Correct exercises with no assistance
		Student satisfies at least one sequence in set (S1-S9) and at least one sequence in set (H5-H11)	*Low autonomy*	Correct exercises with assistance

The graph in Figure 2 shows the distribution of the engagement characteristics of the students under the six identified dimensions. *Persistence*, *Concentration* and *Reflection* are high for most of the students, denoting a fairly high commitment to the task for the majority of the students, whereas *Confidence*, *Autonomy* and *Understanding* show rather variable 563 564 565 566 distributions. This is comprehensible since the level of individual competence and skill can be different, and this influences individual self-confidence and results. *Understanding*, in particular, shows quite significant variations: few students were very quick to solve exercises (*High Understanding*), whereas most of them were able to solve them in a larger interval of time (*Medium Understanding*); the rest of the students were not able to solve the exercise in a predefined interval of time (*Low Understanding*).

Figure 2. Distribution of the engagement characteristics of the students under the engagement dimensions.

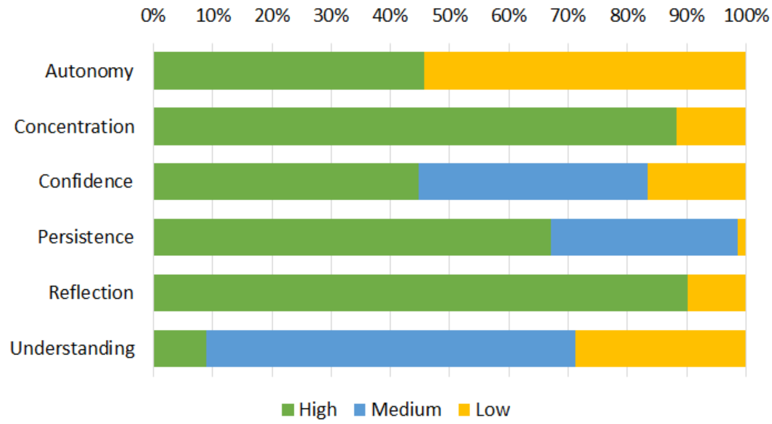
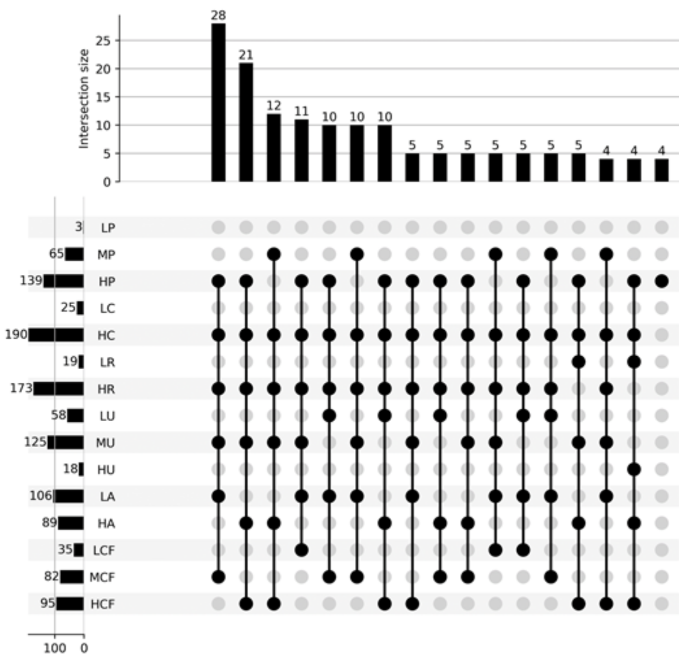


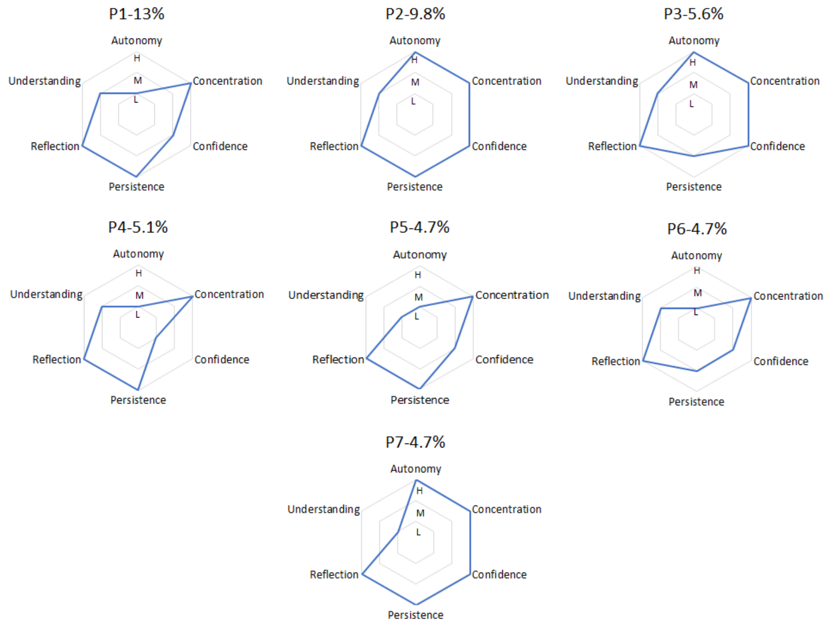
Figure 3 shows the distribution of the students according to the chosen dimensions: each vertical bar represents the number of students who have the same characteristics, which are described by the black dots below (e.g. 28 students have LA=*Low Autonomy*, HU=*High Understanding*, HR=*High Reflection*, MCF=*Medium Confidence*, HC=*High Concentration* and HP=*High Persistence*). The horizontal bars represent the number of students who have that particular characteristics (e.g. 106 students have LA=*Low Autonomy*). The figure shows only the groups composed of at least 4 students.

Figure 3. Distribution of the students according to engagement dimensions and corresponding levels. H=High, L=Low, A=Autonomy, U=Understanding, R=Reflection, CF=Confidence, C=Concentration, P=Persistence



Each student group represents a specific student profile. The radar plots in Figure 4 show the details of the most common profiles. The percentage of students who belong to profile P1, for example, is 13% of the total number of students (215). The considered profiles, together, account for almost 50% of the students (47.4%). Each radar plot shows the level (H=*High*, M=*Medium*, L=*Low*) of the engagement dimensions for the students belonging to a specific profile.

Figure 4. The seven most frequent students' profiles



The takeaways from the student profile distributions presented above are summarized below.

- *Autonomy* and *Confidence* are correlated with each other (see all profiles): either they are both *High*, or they are both *Medium/Low*. This situation makes sense, because *Confidence* is related to students' help request, and *High Confidence* means few help requests), whereas *Autonomy* to correct solutions with or without help (*High Autonomy* means few or no help interventions), and most of the times help requests lead to help interventions.
- In general, all profiles show *High* levels of *Concentration* and of *Reflection*: students are able to stay focused during the whole laboratory session and they are challenged by the proposed exercises.
- Students with profile *P2* show high commitment (*High Persistence* and *High Concentration*), good self-confidence (*High Confidence* and *High Autonomy*) and good results (*Medium Understanding*).
- Students with profile *P7* show high commitment (*High Persistence* and *High Concentration*), good self-confidence (*High Confidence* and *High Autonomy*) but worse results (*Low Understanding*).
- Students with profiles *P1* and *P4* need some help (*Medium/Low Confidence* and *Low Autonomy*) but anyway demonstrate the capability to focus on the task (*High Persistence* and *High Reflection*) and to get good results (*Medium Understanding*).
- Students with profile *P3* and *P6* show some indecision, going back and forth among exercises (*Medium Persistence*), or simply they want to get an overall idea of what they are requested to do in the whole lab session. This behavior does not compromise their performance: they focus on the task (*High Persistence* and *High Reflection*) and

get good results (*Medium Understanding*), with more (profile *P3*) or less (profile *P6*) self-confidence and autonomy.
- Students with profile *P5* show serious difficulties in performing the requested tasks (*Low Understanding*) despite their commitment (*High Concentration* and *High Reflection*) and the help they request and obtain (*Medium Confidence* and *Low Autonomy*).

6.2. Correlation analysis among the engagement dimensions

Here we analyze the pairwise intersections of the six engagement dimensions. Although we considered the pairwise intersections in Figure 5, we show only the most representative. The numbers in the matrices represent the number of students who have the characteristics of the corresponding areas, where *H=High*, *M=Medium*, and *L=Low*.

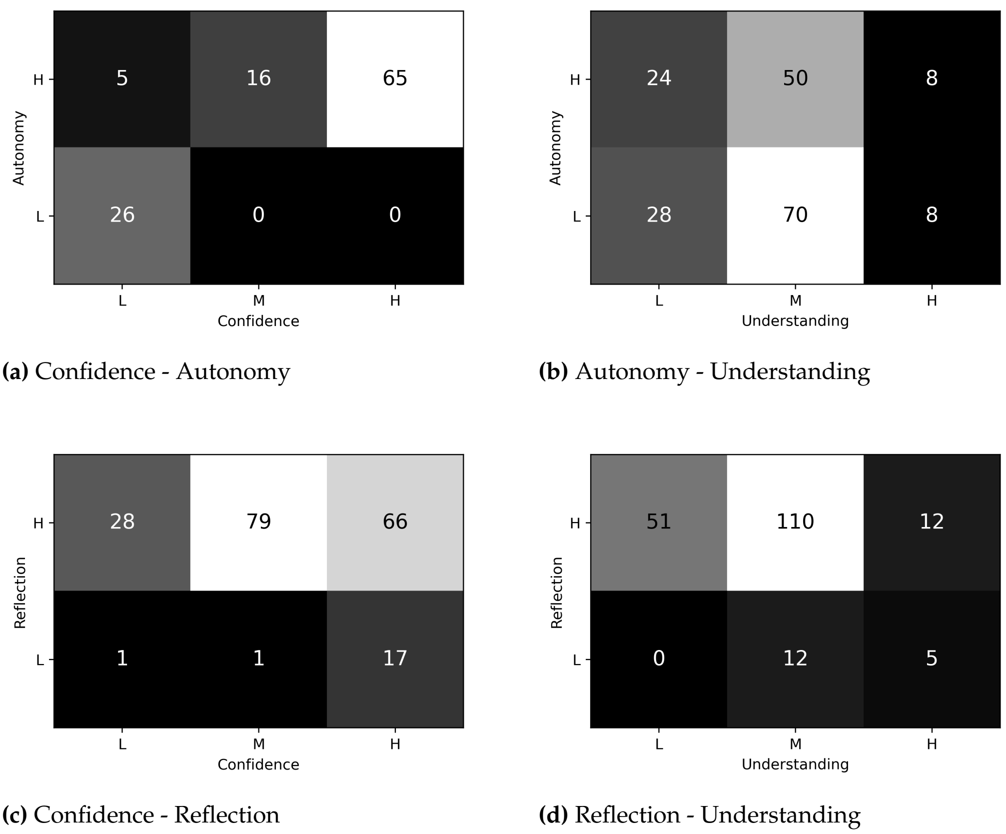
The intersections between *Autonomy* and *Confidence* offer valuable insights into how these two indicators interact. Notably, when *High Autonomy* aligns with *Low Confidence*, we observe a dimensionality of 5. Interestingly, the most substantial intersection occurs when *High Autonomy* combines with *High Confidence*, resulting in a dimensionality of 65. This indicates a strong correlation, implying that individuals with high autonomy levels often coexist with high confidence levels, potentially reinforcing each other.

Conversely, *Autonomy* and *Understanding* are independent. This shows that help interventions, whereas they are generally sufficient to solve the specific task for what they were requested, are not always effective for having a more comprehensive level of understanding, applicable to all the tasks. Besides, they show that the perceived need for external support is very personal and not always related to the actual need.

Similarly, *Reflection* and *Understanding* are not correlated.

Confidence positively influences *Reflection*. Specifically, 66 of students who have *High Confidence* have *High Reflection* too. Conversely, only 26 of the students have *Low Reflection*. This is justifiable because self-confidence helps students to rely on their own capability and to address problems with a reflective approach (compared to a trial-and-error one).

Figure 5. Pairwise intersections among engagement dimensions.



(a) Confidence - Autonomy

(b) Autonomy - Understanding

(c) Confidence - Reflection

(d) Reflection - Understanding

6.3. Discussion

The results shows that the SQL laboratory session involved students who were quite interested and motivated for the whole duration of the session. This is coherent with the fact that laboratory were not compulsory, so students participated because they want to practice and learn, and the lab duration was not excessive (90 minutes).

Students came to the lab session with different backgrounds of competence and skill, depending on the practice they did before the lab. This reflects on the different level of confidence and autonomy demonstrated by the analysis. This background, together with the individual attitude for reflection, influence the understanding dimension, measured in relation to the performance in the assigned task.

We detected some specific student behaviours that were useful for solving the exercise. Specifically, the first one is design by copying and practice that is a common feature in programming, because it is focused on logical thinking rather than on the memorization of the complete code syntax. The second practice is the trial and error schema (also know as "what if"); it reveals the students' attitude of learning from mistakes. It is really common in computer programming learning and it is also typical of gaming thinking. In addition students generally prefer to proceeding step by step, and avoid to skip; however, considering the complexity of some specific exercise (e.g. 5) they risk to be stuck for a long time. We noticed also that most students who participated in the lab have a reflective attitude compared to a trial-and-error one, coherently with what is encouraged during the course.

The analysis of the correlation among the different engagement dimensions considered in the present paper shows that there is a strong link between cognitive and affective engagement, and that that they influence one another. Specifically, *Autonomy* and *Confidence* are strongly correlated, as well as *Confidence* and *Reflection*. A good level of affective engagement reflects on cognitive engagement and vice-versa: self-confidence positively influence the capability to focus effectively on a problem, and in turn good results obviously enhance self-confidence.

The results show also a fairly high correlation between some cognitive engagement dimensions, namely *Concentration*, *Reflection* and *Understanding*: this reflects the steps in which the students face and solve the proposed exercises, focusing on them, reflecting on the possible solutions, and then submitting the answer.

7. Conclusions

This work proposes a method to deeply analyze the student's behaviour during laboratories. It relies on data collected in the context of a B.S. degree course on database design and management. The collected data describe the main activities performed by the participants to a computer lab sess *Confidence* and *Autonomy* are strongly correlated with each other, as shown in diagram (a). Specifically, 68% of students who have *High Confidence* have also *High Autonomy*, whereas 74% of the students who have *Low Confidence* have also *Low Autonomy*. This evidence confirms what previously emerged in the analysis of the most frequent profiles (see section 6.1), and it is explained by the fact that, commonly, when students asked for help (*Confidence*) they received it (*Autonomy*).

Concentration and *Autonomy*, on the opposite side, are independent: 47% of students who have *High Concentration* have *High Autonomy* as well, and 53% have *Low Autonomy*. The general level of *Concentration* is *High* (see Figure 2), but *Autonomy* is a characteristic of the students that is mainly influenced by self-confidence rather than by the capability to focus on a given task.

Autonomy and *Understanding* are also independent, as shown in diagram (b). Specifically, 44% of students who have *High Understanding* have also *High Autonomy* and 44% of them have *Low Autonomy*, while 41% of students who have *Low Understanding* have *High Autonomy* and 48% of them have *Low Autonomy*. This shows that help interventions, whereas they are generally sufficient to solve the specific task for what they were requested, are not always effective for having a more comprehensive level of understanding, applicable to all the tasks. Besides, they show that the perceived need for external support is very personal and not always related to the actual need.

Most students have *High Concentration* and *High Reflection* (as shown in Figure 2), and they are correlated with each other: 87% of students have *High Concentration* have *High Reflection* too, and only 9% of them have *Low Reflection* as well. This is understandable, because the capability to focus on a task influences the attitude to apply a more reflective approach in problem solving.

Confidence positively influences *Reflection*, as shown in diagram (c). Specifically, 69% of students who have *High Confidence* have *High Reflection* too. Conversely, only 18% of the students have *Low Reflection*, and 84% of students who have *High* or *Medium Confidence* have also *High Reflection*. This is justifiable because self-confidence helps students to rely on their own capability and to address problems with a reflective approach (compared to a trial-and-error one).

The implication that *Reflection* positively influence *Understanding* clearly emerges from the performed analyses, as shown in diagram (d). Specifically, 71% of students with *High Reflection* have *High* or *Medium Understanding* whereas only 29% have *Low Understanding*, and only 28% of students who have *High Understanding* have *Low Reflection*. The attitude to face problem in a more reflective way has a positive influence to apply what has been learned in the following ones. The sequence of exercises was proposed by the teacher with this goal in mind, to progressively build competence and skills in the specific subject.

No specific correlation was found between *Persistence* and the other dimensions, possibly because the persistence level was high for almost all the students: the laboratory was not compulsory so the participating students were mainly committed to it, with a good level of behavioural engagement. If the laboratory would be compulsory, probably the results would have been different, with a variable level of behavioural engagement that could have influenced cognitive and affective engagement aspects.

7.1. Discussion

The results shows that the SQL laboratory session involved students who were quite interested and motivated for the whole duration of the session. This is coherent with the fact that laboratory were not compulsory, so students participated because they want to practice and learn, and the lab duration was not excessive (90 minutes).

Students came to the lab session with different backgrounds of competence and skill, depending on the practice they did before the lab. This reflects on the different level of confidence and autonomy demonstrated by the analysis. This background, together with the individual attitude for reflection, influence the understanding dimension, measured in relation to the performance in the assigned task.

We detected some specific student behaviours that were useful for solving the exercise. Specifically, the first one is design by copying and practice that is a common feature in programming, because it is focused on logical thinking rather than on the memorization of the complete code syntax. The second practice is the trial and error schema (also know as "what if"); it reveals the students' attitude of learning from mistakes. It is really common in computer programming learning and it is also typical of gaming thinking.ion on the SQL language. The experiment considered various types of events such as the accesses to exercises, the correct answers submissions, the errors, the assistance requests and the teaching assistants' interventions.

The paper explores the use of sequential pattern mining techniques to analyze the temporal correlations between the student-related events occurred during the lab sessions. Based on the extracted patterns, students are profiled according to their levels of engagement in various dimensions. By examining the most significant extracted patterns and profiles, it was possible to get a detailed view of the students' activities. This allowed us to recognize cause-effect correlations, positive aspects and points of criticism in order to improve the lab experience.

The pattern extraction phase allowed us to define a number of engagement key indicators that are useful for assessing the level of behavioural, cognitive and affective engagement of the students during the computer lab. The students demonstrated a very good level of behavioural engagement (*Persistence*), a satisfactory level of cognitive engagement (*Concentration*, *Reflection*, *Understanding* and *Autonomy*), where *Autonomy* and *Understanding* are the most variable dimensions, being dependent on the individual background of competence and skills. About the level of affective engagement (*Confidence*), it is highly variable, depending on the individual capability to face the proposed tasks. Besides, the engagement analysis highlighted some interesting correlations among the identified engagement dimensions. The latter findings, in particular, showed that the cognitive dimensions of engagement are strictly correlated with the affective dimensions, and that they positively influence one another.

Future works will focus on tracing, collecting, and analyzing students' data in laboratories related to different courses. The key goal is to discover which patterns are universal and which ones are subject-dependent. We will also explore the use of different learning environments (both online and in presence) and the application of a similar approach to event sequence mining to data acquired in different learning contexts, such as persuasive and recruitment games.

Author Contributions: Conceptualization, L.C., L.C., and L.F.; methodology, L.C., L.C., and L.F.; software, L.C. and L.C.; validation, L.C., L.C., and L.F.; formal analysis, L.C., L.C., and L.F.; investigation, L.C., L.C., and L.F; resources, L.C., L.C., and L.F.; data curation, L.C., L.C., and L.F.; writing—original draft preparation, L.C., L.C., and L.F.; writing—review and editing, L.C., L.C., and L.F.; visualization, L.C., L.C., and L.F.; supervision, L.C. and L.F. All authors have read and agreed to the published version of the manuscript.

Funding: his research received no external funding

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: The data presented in this study are available on request from the corresponding author

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bedenlier, S.; Bond, M.; Buntins, K.; Zawacki-Richter, O.; Kerres, M., Learning by Doing? Reflections on Conducting a Systematic Review in the Field of Educational Technology. In *Systematic Reviews in Educational Research: Methodology, Perspectives and Application*; Zawacki-Richter, O.; Kerres, M.; Bedenlier, S.; Bond, M.; Buntins, K., Eds.; Springer Fachmedien Wiesbaden: Wiesbaden, 2020; pp. 111–127. https://doi.org/10.1007/978-3-658-27602-7_7.
2. Heradio, R.; de la Torre, L.; Galan, D.; Cabrerizo, F.J.; Herrera-Viedma, E.; Dormido, S. Virtual and Remote Labs in Education. *Comput. Educ.* **2016**, *98*, 14–38. https://doi.org/10.1016/j.compedu.2016.03.010.
3. Romero, C.; Ventura, S. Educational data mining and learning analytics: An updated survey. *WIREs Data Mining and Knowledge Discovery* **2020**, *10*, e1355, [https://onlinelibrary.wiley.com/doi/pdf/10.1002/widm.1355]. https://doi.org/10.1002/widm.1355.
4. Bond, M.; Buntins, K.; Bedenlier, S.; Zawacki-Richter, O.; Kerres, M. Mapping research in student engagement and educational technology in higher education: a systematic evidence map. *International Journal of Educational Technology in Higher Education* **2020**, *17*. https://doi.org/10.1186/s41239-019-0176-8.
5. Marks, H. Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal* **2000**, *37*, 153–184. https://doi.org/10.3102/00028312037001153.
6. Fredricks, J.A.; Blumenfeld, P.C.; Paris, A.H. School engagement: Potential of the concept, state of the evidence. *Review of Educational Research* **2004**, *74*, 59–109. https://doi.org/10.3102/00346543074001059.
7. Eccles, J. Engagement: Where to next? *Learning and Instruction* **2016**, *43*, 71–75.
8. Akçapinar, G.; Chen, M.R.A.; Majumdar, R.; Flanagan, B.; Ogata, H. Exploring Student Approaches to Learning through Sequence Analysis of Reading Logs. In Proceedings of the Proceedings of the Tenth International Conference on Learning Analytics & Knowledge; Association for Computing Machinery: New York, NY, USA, 2020; LAK '20, p. 106–111. https://doi.org/10.1145/3375462.3375492.
9. Sun, Y.; Guo, Y.; Zhao, Y. Understanding the determinants of learner engagement in MOOCs: An adaptive structuration perspective. *Computers & Education* **2020**, *157*, 103963. https://doi.org/https://doi.org/10.1016/j.compedu.2020.103963.

10. Xie, K.; Heddy, B.C.; Greene, B.A. Affordances of using mobile technology to support experience-sampling method in examining college students' engagement. *Computers & Education* **2019**, *128*, 183 – 198. https://doi.org/https://doi.org/10.1016/j.compedu.2018.09.020.
11. Sunar, A.S.; White, S.; Abdullah, N.A.; Davis, H.C. How Learners' Interactions Sustain Engagement: A MOOC Case Study. *IEEE Transactions on Learning Technologies* **2017**, *10*, 475–487.
12. Yousuf, B.; Conlan, O. Supporting Student Engagement Through Explorable Visual Narratives. *IEEE Transactions on Learning Technologies* **2018**, *11*, 307–320.
13. Bergdahl, N.; Nouri, J.; Fors, U.; Knutsson, O. Engagement, disengagement and performance when learning with technologies in upper secondary school. *Computers & Education* **2020**, *149*, 103783. https://doi.org/https://doi.org/10.1016/j.compedu.2019.103783.
14. Taipalus, T.; Seppänen, V. SQL Education: A Systematic Mapping Study and Future Research Agenda. *ACM Trans. Comput. Educ.* **2020**, *20*. https://doi.org/10.1145/3398377.
15. Zaki, M.J. SPADE: An Efficient Algorithm for Mining Frequent Sequences; Kluwer Academic Publishers: USA, 2001; Vol. 42, p. 31–60.
16. Balamuralithara, B.; Woods, P.C. Virtual laboratories in engineering education: The simulation lab and remote lab. *Computer Applications in Engineering Education* **2009**, *17*, 108–118, [https://onlinelibrary.wiley.com/doi/pdf/10.1002/cae.20186]. https://doi.org/10.1002/cae.20186.
17. Parker, J.; Cupper, R.; Kelemen, C.; Molnar, D.; Scragg, G. Laboratories in the Computer Science Curriculum. Routledge, 1990, Vol. 1, pp. 205–221. https://doi.org/10.1080/0899340900010303.
18. Knox, D.; Wolz, U.; Joyce, D.; Koffman, E.; Krone, J.; Laribi, A.; Myers, J.P.; Proulx, V.K.; Reek, K.A. Use of Laboratories in Computer Science Education: Guidelines for Good Practice: Report of the Working Group on Computing Laboratories. In Proceedings of the Proceedings of the 1st Conference on Integrating Technology into Computer Science Education; Association for Computing Machinery: New York, NY, USA, 1996; ITiCSE '96, p. 167–181. https://doi.org/10.1145/237466.237644.
19. Aversano, L.; Canfora, G.; De Lucia, A.; Stefanucci, S. Understanding SQL through iconic interfaces. In Proceedings of the Proceedings 26th Annual International Computer Software and Applications, 2002, pp. 703–708.
20. Sadiq, S.; Orlowska, M.; Sadiq, W.; Lin, J. SQLator: An Online SQL Learning Workbench. In Proceedings of the Proceedings of the 9th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education; Association for Computing Machinery: New York, NY, USA, 2004; ITiCSE '04, p. 223–227. https://doi.org/10.1145/1007996.1008055.
21. Mitrovic, A. Learning SQL with a computerized tutor. 1998, Vol. 30, pp. 307–311. https://doi.org/10.1145/273133.274318.
22. Romero, C.; Ventura, S. Guest Editorial: Special Issue on Early Prediction and Supporting of Learning Performance. *IEEE Transactions on Learning Technologies* **2019**, *12*, 145–147.
23. Ahadi, A.; Behbood, V.; Vihavainen, A.; Prior, J.; Lister, R. Students' Syntactic Mistakes in Writing Seven Different Types of SQL Queries and Its Application to Predicting Students' Success. In Proceedings of the Proceedings of the 47th ACM Technical Symposium on Computing Science Education; Association for Computing Machinery: New York, NY, USA, 2016; SIGCSE '16, p. 401–406. https://doi.org/10.1145/2839509.2844640.
24. Figueira, A.S.; Lino, A.S.; Paulo, S.S.; Santos, C.A.M.; Brasileiro, T.S.A.; Del Pino Lino, A. Educational data mining to track students performance on teaching learning environment LabSQL. In Proceedings of the 2015 10th Iberian Conference on Information Systems and Technologies (CISTI), 2015, pp. 1–6.
25. Zaldivar, V.; Pardo, A.; Burgos, D.; Delgado-Kloos, C. Monitoring student progress using virtual appliances: A case study. *Computers & Education* **2012**, *58*, 1058–1067.
26. Lertnattee, V.; Pamonsinlapatham, P. Blended Learning for Improving Flexibility of Learning Structure Query Language (SQL). In Proceedings of the Blended Learning. New Challenges and Innovative Practices; Cheung, S.K.; Kwok, L.f.; Ma, W.W.; Lee, L.K.; Yang, H., Eds.; Springer International Publishing: Cham, 2017; pp. 343–353.
27. Fang, A.D.; Chen, G.L.; Cai, Z.R.; Cui, L.; Harn, L. Research on Blending Learning Flipped Class Model in Colleges and Universities Based on Computational Thinking—"Database Principles" for Example. *Eurasia Journal of Mathematics, Science and Technology Education* **2017**, *13*, 5747–5755.
28. Soflano, M.; Connolly, T.M.; Hainey, T. An application of adaptive games-based learning based on learning style to teach SQL. *Computers & Education* **2015**, *86*, 192 – 211. https://doi.org/https://doi.org/10.1016/j.compedu.2015.03.015.
29. Dol, S.M. Use of Self-Created Videos for Teaching Structured Query Language (SQL) using Flipped Classroom Activity. *Journal of Engineering Education Transformations* **2020**, *33*.
30. Tsai, Y.S.; Gasevic, D. Learning Analytics in Higher Education — Challenges and Policies: A Review of Eight Learning Analytics Policies. In Proceedings of the Proceedings of the Seventh International Learning Analytics & Knowledge Conference; Association for Computing Machinery: New York, NY, USA, 2017; LAK '17, p. 233–242. https://doi.org/10.1145/3027385.3027400.
31. Zaki, M.J.; Meira, Jr, W. *Data Mining and Machine Learning: Fundamental Concepts and Algorithms*, 2 ed.; Cambridge University Press, 2020. https://doi.org/10.1017/9781108564175.
32. Wong, J.; Khalil, M.; Baars, M.; de Koning, B.B.; Paas, F. Exploring sequences of learner activities in relation to self-regulated learning in a massive open online course. 2019, Vol. 140, p. 103595.
33. Appleton, J.J.; Christenson, S.L.; Furlong, M.J. Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools* **2008**, *45*, 369–386. https://doi.org/10.1002/pits.20303.

34. Kuh, G.D.; Cruce, T.M.; Shoup, R.; Kinzie, J.; Gonyea, R.M. Unmasking the effects of student engagement on first-year college grades and persistence. *The Journal of Higher Education* **2008**, *79*, 540–563. https://doi.org/10.1080/00221546.2008.11772116.
35. Henrie, C.R.; Halverson, L.R.; Graham, C.R. Measuring Student Engagement in Technology-Mediated Learning. *Comput. Educ.* **2015**, *90*, 36–53. https://doi.org/10.1016/j.compedu.2015.09.005.
36. Martin, A.J. Motivation and engagement: Conceptual, operational, and empirical clarity. In *Handbook of research on student engagement*; Christenson, S.L.; Reschly, A.L.; Wylie, C., Eds.; Springer: Boston, MA, USA, 2012; pp. 303–311. https://doi.org/10.1007/978-1-4614-2018-7_14.
37. Redmond, P.; Heffernan, A.; Abawi, L.; Brown, A.; Henderson, R. An Online Engagement Framework for Higher Education. *Online Learning* **2018**, *22*. https://doi.org/10.24059/olj.v22i1.1175.
38. Reeve, J. A self-determination theory perspective on student engagement. In *Handbook of research on student engagement*; Christenson, S.L.; Reschly, A.L.; Wylie, C., Eds.; Springer: Boston, MA, USA, 2012; pp. 149–172. https://doi.org/10.1007/978-1-4614-2018-7_7.
39. Skinner, E.; Pitzer, J.R. Developmental dynamics of student engagement, coping, and everyday resilience. In *Handbook of research on student engagement*; Christenson, S.L.; Reschly, A.L.; Wylie, C., Eds.; Springer: Boston, MA, USA, 2012; pp. 21–44. https://doi.org/10.1007/978-1-4614-2018-7_2.
40. Lee, J.; Song, H.; Hong, A. Exploring Factors, and Indicators for Measuring Students' Sustainable Engagement in e-Learning. *Sustainability* **2008**, *11*. https://doi.org/10.3390/su11040985.
41. Zepke, N. Student engagement in neo-liberal times: What is missing? *Higher Education Research and Development* **2018**, *37*, 433–446. https://doi.org/10.1080/07294360.2017.1370440.
42. Smith, T.; Lambert, R. A systematic review investigating the use of twitter and Facebook in university-based healthcare education. *Health Education* **2014**, *114*, 347–366. https://doi.org/10.1108/HE-07-2013-0030.
43. Kahu, E.R. Framing student engagement in higher education. *Studies in Higher Education* **2013**, *38*, 758–773. https://doi.org/10.1080/03075079.2011.598505.
44. Jung, Y.; Lee, J. Learning engagement and persistence in massive open online courses (MOOCS). *Computers & Education* **2018**, *122*, 9–22. https://doi.org/10.1016/j.compedu.2018.02.013.
45. Loney, K. *Oracle Database 11g The Complete Reference*, 1 ed.; McGraw-Hill, Inc.: USA, 2008.
46. Pei, J.; Han, J.; Mortazavi-Asl, B.; Pinto, H.; Chen, Q.; Dayal, U.; Hsu, M. PrefixSpan: Mining Sequential Patterns by Prefix-Projected Growth. In Proceedings of the Proceedings of the 17th International Conference on Data Engineering, April 2-6, 2001, Heidelberg, Germany; Georgakopoulos, D.; Buchmann, A., Eds. IEEE Computer Society, 2001, pp. 215–224. https://doi.org/10.1109/ICDE.2001.914830.
47. Zaki, M.J. Sequence Mining in Categorical Domains: Incorporating Constraints. In Proceedings of the Proceedings of the Ninth International Conference on Information and Knowledge Management; Association for Computing Machinery: New York, NY, USA, 2000; CIKM '00, p. 422–429. https://doi.org/10.1145/354756.354849.
48. Ho, J.W.K.; Lukov, L.; Chawla, S. Sequential Pattern Mining with Constraints on Large Protein Databases. In Proceedings of the International Conference on Management of Data, 2005.
49. Berge, O.; Borge, R.; Fjuk, A.; Kaasbøll, J.; Samuelsen, T. Learning Object-Oriented Programming. 2003.
50. Berglund, A.; Eckerdal, A. Learning Practice and Theory in Programming Education: Students' Lived Experience. 2015, pp. 180–186. https://doi.org/10.1109/LaTiCE.2015.49.
51. Cagliero, L.; De Russis, L.; Farinetti, L.; Montanaro, T. Improving the Effectiveness of SQL Learning Practice: A Data-Driven Approach. In Proceedings of the 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), 2018, Vol. 01, pp. 980–989.