Using Educational Data From Teaching And Learning To Inform Teachers' Reflective Educational Design In Inquiry-Based STEM Education

Stylianos Sergis[1], Demetrios G. Sampson*[1,2], María Jesús Rodríguez-Triana[3,4], Denis Gillet[3], Lina Pelliccione[2], Ton de Jong[5]

[1] Department of Digital Systems, University of Piraeus, Piraeus, GR-18532, Greece. E-mail: steliossergis@gmail.com
[2] School of Education, Curtin University, Bentley Campus, Building 501.Level 4, Perth, Australia. E-mails: Demetrios.Sampson@curtin.edu.au, L.Pelliccione@curtin.edu.au
[3] REACT, School of Engineering, École Polytechnique Fédérale de Lausanne, Switzerland, E-mails: maria.rodrigueztriana@epfl.ch, denis.gillet@epfl.ch
[4] Center of Excellence in Educational Innovation, School of Digital Technologies, Tallinn University, Estonia, E-mail: mjrt@tlu.ee
[5] Department of Instructional Technology, Faculty of Behavioral Sciences, University of Twente, 7500AE Enschede, Netherlands, E-mail: a.j.m.dejong@utwente.nl
* Corresponding author

Highlights

- Tool-supported Guidance is essential for effective Inquiry-based education
- Teaching and Learning Analytics (TLA) can support teachers provide appropriate Guidance
- The TLA method and supporting tool provides analyses of the level of Guidance in Inquiry-based scenarios
- Analyses of the design can be investigated against customizable learners' data and profiles
- Insights from these combined analyses could help teachers improve their teaching designs

Abstract: Science, Technology, Engineering and Mathematics (STEM) education is recognized as a top priority for school education worldwide and Inquiry-based teaching and learning is identified as one of the most dominant approaches. To effectively engage individual students in Inquiry tasks, appropriate guidance needs to be provided, usually by combining different digital tools such online labs, data analysis tools and modelling tools. This is a cumbersome task for teachers to perform manually since it involves (a) assessing during the education design, the type and level of tool-supported guidance to be provided to students and (b) potentially refining this level and types to meet the guidance needs of individual students based on educational data from the delivery of the educational design. Thus, in our research we target to investigate how to support this process with educational data analytics methods and tools from both the design and the delivery of educational designs, that inform teachers' decision making for systematic reflection. To this end, the contribution of this paper is the design and evaluation of a novel "*Teaching and Learning*" Analytics method and supporting research prototype tool, extending the scope of purely learning analytics methods, to (a) analyze inquiry-based educational designs in terms of the tool-supported guidance they offer and (b) relate these analyses to students' educational data that are already being collected by existing learning analytics systems, so as to increase teachers' awareness and understanding and scaffold their reflection. A two-layer evaluation methodology was adopted to evaluate both the capacity of our method to analyze educational designs in terms of appropriate guidance as well as to investigate whether the insights generated by the method offer statistically significant indicators that impact students' activity during the delivery of these educational designs. The results obtained, based on real-life educational data, argue that the proposed method and tool can support teachers to accurately analyse Inquiry-based educational designs and receive meaningful insights to improve and tailor students' learning experiences. The insights of this work aim to contribute in the research field of cognitive data analytics for teaching and learning, by investigating new ways to combine analyses of the educational design and the students' activity, so as to inform teachers' reflective decision making from a holistic perspective.

Keywords: Educational Data Analytics; Teaching and Learning Analytics; Teaching Analytics; Learning Analytics; Guidance; Inquiry-based teaching; STEM Education

1. Introduction

Science, Technology, Engineering and Mathematics (STEM) has been recognized as a top priority for school education worldwide (Johnson et al., 2015). It posits the standpoint that student-centred teaching approaches, such as the inquiry-based approach, should be exploited for cultivating (among others) students' inquiry skills through engagement with diverse challenging tasks (Freeman et al., 2015), for example formulating research hypotheses, planning and conducting experiments, as well as analysing and interpreting data.

There is a significant body of evidence on the benefits of inquiry-based learning on students (e.g., enhanced cognitive outcomes (Donnelly et al., 2014) and motivation (Huang et al., 2013), therefore, explicit focus is increasingly being placed to support students to reap these benefits. However, Inquiry tasks are often challenging for students, who are usually in need of

guidance (d' Angelo et al., 2013) to successfully engage with and complete them. Emerging technologies are commonly employed to effectively provide this supporting guidance.

A prime example of technologies used in this regard are *online labs.* Online labs, which include virtual and remote labs (Zervas et al., 2015) are utilized in STEM **Inquiry-based Educational Designs (IED)** given their affordances to foster more 'hands-on' learning experiences for students (de Jong et al., 2014) by operating virtual or physical equipment as well as their capacity to provide guidance to support students' engagement in the Inquiry tasks (d' Angelo et al., 2013). However, often online labs provide only a subset of the required guidance needed to facilitate students perform the required tasks throughout the Inquiry cycle (de Jong et al., 2013). So, in this case, appropriate external tools (from now on: *guidance tools*) may need to be selected to complement these design

shortcomings and offer technology-enhanced guidance (d' Angelo et al., 2013).

This can create additional, and potentially cumbersome, tasks for the teachers including how to effectively analyze their IED in terms of the level and appropriateness of the provided guidance, and also how to meaningfully improve it, by also taking into account evidence from students' learning data. In this context, purely Learning Analytics methods tend to focus on the latter, while marginally (if at all) consider the aspect of how the educational design might affect the student activity. Therefore, the main standpoint of the present work is that these reflective processes can be supported by exploiting the potential of emerging "**Teaching and Learning" Analytics (TLA)** methods and tools (Sergis & Sampson, 2017). The reason for this standpoint is that TLA methods have been increasingly used to support teachers' *systematic reflection* on their teaching practice, extending the scope of purely Learning Analytics methods, using educational data collected from students (and/or themselves), in a process commonly termed *teacher inquiry*[1] (Avramides et al., 2015).

Therefore, the present work capitalizing on prior work (Sergis & Sampson, 2016a) proposes a TLA method to support teachers' systematic reflection on their IED in terms of tool-supported guidance. The TLA method is also incorporated within a research prototype tool that hosts the required functionalities of the proposed method [2] (discussed further in Section 3). From the one hand, the proposed method enables teachers to analyse and self-assess the level and types of tool-supported guidance they have provided before delivery, allowing for early-on (re)design. On the other hand, (which refers to the main novelty of this paper compared to the prior work), the proposed method has been extended to receive existing educational data collected from Learning Analytics tools of digital learning environments, through .json files (this work focuses on the 'Graasp' learning environment (Rodriguez-Triana et al., 2014), as a specific case study). By capitalizing on these Learning Analytics tools and the educational data they collect, the tool can build student profiles which are directly related to the analysis of the IED that students engaged with.

In this way, the proposed method provides a new context to exploit student data already being collected by Learning Analytics tools, by allowing their joint processing with data from the educational design. This emerging field of *Teaching and Learning* Analytics offers the capacity to elicit and visualize the guidance types and level provided in the educational design (Teaching Analytics) and, also, investigate the impact of this guidance on the activity and performance of individual (or groups of) students (collected via Learning Analytics tools) so as to inform teachers' decision making on designing more tailored IEDs.

The paper also presents the results of a two-layer evaluation protocol to thoroughly investigate the accuracy and potential of the proposed method for informing teachers' reflection. The **first layer of evaluation** aimed to investigate whether the proposed TLA method can correctly analyze existing IED in terms of tool-provided guidance. A preliminary version of this layer was reported in Sergis & Sampson (2016a). However, this paper reports on an enhanced evaluation of this layer, comprising a more elaborate set of metrics focusing on classification precision and accuracy. The **second layer of evaluation**, introduced for the first time in this work, aimed to investigate whether the analyses generated by the tool provide statistically significant indicators that can impact diverse students' types of activity in the delivery of IED, and can, therefore, provide meaningful insights for (re)design.

The evaluation protocol focuses on a case study, so as to provide a robust research proof-of-concept. In particular, the educational data and IED utilized in the evaluations were collected in the context of the major European project "Global Online Science Labs for Inquiry Learning at School" (Go-Lab) (http://www.go-lab-project.eu), which employed the Graasp digital learning environment for collecting students' educational data from the delivery of IED.

The remainder of the paper is structured as follows: Section 2 presents the background of this work, namely the concept of guidance in technology-enhanced Inquiry-based STEM Education. Furthermore, it outlines the potential of TLA to address existing challenges, presenting the specific contribution of this work. Section 3 presents the proposed TLA tool that hosts the proposed method, in terms of components and types of insights provided to the teacher. Section 4 outlines the two-layer evaluation methodology. Section 5 presents the evaluation results obtained. Finally, section 6 discusses conclusions and outlines potential future work.

2. Theoretical Background and Literature Review

2.1 Guidance in technology-enhanced Inquiry-based STEM Education

The Inquiry-based approach in STEM education adopts an exploratory standpoint, promoting students' active participation and self-regulated discovery of knowledge (Sharples et al., 2015). More specifically, it posits the notion that students should be engaged in a cycle of Inquiry phases (and tasks) to create their own reasoning on real-life phenomena (Pedaste et al., 2015). In a recent state-of-the-art review, this cycle of Inquiry phases was defined as follows, comprising specific *tasks* (Pedaste et al., 2015)[3]:

- **Phase 1: Orientation:** to *Stimulate students' curiosity on the phenomenon* to be studied through provision of information
- **Phase 2: Conceptualization** of the phenomenon: to guide students *Formulate hypotheses* to investigate and *Plan an appropriate experimentation process*,
- **Phase 3: Investigation** of the phenomenon: to *Perform experimentation process* and *Collect and analyse experimentation data*
- **Phase 4: Conclusion** on the investigation: to *Draw conclusions from and Reflect on* the experimentation results;
- **Phase 5: Discussion**: to *Communicate findings* from the previous phases.

The Inquiry-based approach has received a strong attention globally, based on the reported benefits it can deliver (among

---

[1] To avoid misconceptions, 'inquiry' related to teachers' investigation on their practice will be written in lowercase, whereas 'Inquiry' related to the teaching approach and students' skills will be capitalized.

[2] It is noted that at the current point, this research prototype tool includes all functionalities, however it is not yet available for end-user exploitation, pending front-end development and evaluation with teachers.

[3] This conceptualization has been adopted in this work

others) on students' attainment of educational objectives related to the subject domain (Donnelly et al., 2014), level of motivation (Hwang et al., 2013) as well as the cultivation of problem-solving and Inquiry skills (Gillies et al., 2014). Furthermore, as recent evidence indicates, these benefits are further enhanced when the Inquiry process is supported by online labs (d' Angelo et al., 2013), offering students unique learning experiences (de Jong et al., 2013).

However, to reap the aforementioned benefits, students are in need of **guidance** during the technology-supported Inquiry cycle, so as to effectively perform the usually challenging tasks (Wu & Pedersen, 2011). Such guidance is commonly supported by tools (such as online labs) which can be planned in advance of the IED delivery, so as to be available to the students throughout the learning process (Belland, 2017).

The concept of Guidance has been classified in a set of **types**, as follows (de Jong & Lazonder, 2014):
- **Process constraints,** which limit the options (e.g., variables) that students should consider in the Inquiry process.
- **Performance dashboards**, which provide information on the student's performance and/or progress.
- **Prompts,** which provide (usually) text-based hints to the students on what to do in each Inquiry task.
- **Heuristics**, which provide (usually) text-based hints that specifically present instructions to be followed.
- **Direct presentation of information**, which provides background information to ensure that students have the necessary knowledge to engage in a task.
- **Scaffolds,** which facilitate the students in conducting specific tasks. For example, scaffolds can include modeling tools and data analysis tools.

The added value of providing these types of tool-supported guidance to students during Inquiry has been supported by a significant body of research (Quintana et al., 2004). For example, (Kim & Hannafin, 2011) showed that technology-enhanced guidance assisted students in effectively reflecting and communicating their understandings, as well as engaging in tasks related to data logging and management of procedures. Other reported benefits of guidance in IED include support for modeling of their hypotheses (Gijlers & de Jong, 2009), and analyzing scientific data (Eckhardt et al., 2013).

Based on the above, it is reasonable to argue that providing appropriate tool-supported guidance to students is an essential part of effectively assisting them to engage in the previously mentioned Inquiry tasks commonly designed in IED. For example, a *scaffolding* tool can be used for assisting students to *formulate hypotheses* whereas *prompts* could be employed to provide students with hints on *drawing conclusions from and reflect on* their experimentation findings. Such combinations of guidance types with Inquiry tasks will be referred to, in this work, as **Guidance Instances**.

In this context, a common issue in technology-supported Inquiry teaching and learning is that online labs (which are at the core of the IED) usually provide only a subset of the Guidance Instances required throughout the Inquiry cycle (de Jong et al., 2013; Gillet et al., 2013). Therefore, in these cases, external *guidance tools*, such as modeling, hypothesis formulation and data analysis tools, need to be utilized to complement these guidance requirements. This issue creates additional tasks to STEM teachers: to provide adequate guidance in their IED by effectively combining different guidance tools **(Task #1)** and to refine their IED so as to accommodate more appropriate guidance to students or groups **(Task #2)**.

The standpoint of this work is that to effectively address both these tasks, the potential of the emerging Teaching and Learning Analytics (TLA) could be exploited. The following section discusses the state-of-the-art in TLA in terms of addressing these tasks and outlines the specific contribution of this work for progressing this state-of-the-art.

2.2 Teaching and Learning Analytics in Inquiry-based STEM education

TLA introduces a synergy of the potential of existing Analytics strands, as follows:
- the potential of *Teaching Analytics* to analyze the educational designs, e.g., in the constituent elements (such as learning and assessment activities and educational resources/tools) as well as the potential interrelations between these elements and the learning experiences they can foster (e.g., Sergis & Sampson, 2016a);
- the potential of *Learning Analytics* to support the "measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs" (Siemens & Long, 2011), aiming to help teachers provide scaffolding and support to individual (or groups of) students during the learning process.

More specifically, TLA argues that insights generated by Learning Analytics methods and tools can be jointly processed with the analyzed (through Teaching Analytics) elements of teaching practice, and therefore support teachers to improve their educational design and delivery in a more systematic way (Lockyer et al., 2013). In this way, the concept of TLA is highly aligned to the concept of *teacher inquiry* (Sergis & Sampson, 2017). More specifically, teacher inquiry can be defined as a process in which "*teachers identify questions for investigation in their practice and then design a process for collecting evidence about student learning that informs their subsequent educational designs*" (Avramides et al., 2015). Indeed, this evidence-based teacher reflection has forged a highly emerging research field not only in STEM education but the wider technology-enhanced education as well (Haya et al., 2015; Wasson et al., 2016).

Focusing on STEM education, there have been initial approaches to use TLA for addressing the two aforementioned teachers' tasks. Regarding **Task #1** (primarily related to Teaching Analytics**)**, any initial work has been conducted in the form of decision support within existing IED *authoring tools*, i.e., Teaching Analytics functionalities to support teachers in evaluating the tool-supported guidance they have designed for their students. To outline this initial work, a review of widely-known IED authoring tools was performed in order to identify the level of (automated) support they provided to teachers regarding analyzing the IED and highlighting potential oversights regarding appropriate guidance tools. Table 1 depicts the outcomes of this review, namely a description of the 'affordances' to facilitate teachers' design for and assess the

level (and appropriateness/completeness) of tool-supported guidance provided.

As Table 1 depicts, even though the majority of existing IED authoring tools provide functionalities to design for Guidance, they do not explicitly afford functionalities for allowing teachers assess the level and appropriateness (or completeness) of the Guidance tools they have included in their IED. Therefore, it is reasonable to assume that Task #1 is currently under-supported by existing approaches, placing this cumbersome process at the hands of teachers.

Table 1: Overview of IED authoring tools

IED Authoring Tool	Ref.	Designing for Guidance provision
SCY	De Jong et al., 2010	Customizable types of guidance available to be included in IED – No feedback on guidance appropriateness/ completeness
STOCHASMOS	Kyza et al., 2011	Existing types of guidance available to be included in IED– No feedback on guidance appropriateness/ completeness
Co-Lab	van Joolingen et al., 2005	Existing types of guidance available to be included in IED – No feedback on guidance appropriateness/ completeness
nQuire	Mulholland et al., 2012	No explicit consideration for guidance tools
WISE	Slotta, 2004	Existing types of guidance available to be included in IED – No feedback on guidance appropriateness/ completeness
Inspiring Science Education	Zervas &. Sampson, In Press	External guidance tools accommodated – No feedback on guidance appropriateness/ completeness
WeSpot	Mikroyannidis et al., 2013	Existing types of guidance available to be included in IED – No feedback on guidance appropriateness/ completeness
Inquiry Island	White et al., 2002	Customizable types of guidance available to be included in IED – No feedback on guidance appropriateness/ completeness
Graasp	Vozniuk et al., 2016	External guidance tools available (and recommended) to be included in IED – No feedback on guidance appropriateness / completeness

Regarding **Task #2** (relevant to TLA), this work capitalized on a recent systematic review of TLA research literature (Sergis & Sampson, 2017). One of the findings of this work was that existing TLA methods (both generic as well as STEM-specific) provide a common pattern of insights to support teacher inquiry. In particular, in terms of *analysis of educational design*, existing methods focus on outlining the elements that comprise it (i.e., the learning/assessment activities and the educational resources/tools used to support the activities). Building on this analysis of educational design, the *insights* offered primarily targets to inform the teacher on the extent that students interacted with each of these elements (e.g., frequency and patterns of access) (e.g., Ali et al., 2012; Dyckhoff et al., 2012). This pattern is also evident in the small set of pioneering TLA tools in the context of STEM Education.

For example, Vozniuk et al. (2015) presented a method to support teachers' reflection on their teaching practice by providing insights of students' activity, mainly focused on the time spent on each Inquiry phase and outlining which learning activities were accessed by the students and in which sequence (access patterns). Monroy et al. (2014) presented a method and tool that aims to help teachers improve their practice, based on teacher activity data from the delivery of IED (e.g., Inquiry

phase access patterns, time spent online and on each Inquiry phase of IEDs) and can be used to support teachers' inquiry on both their IED as well as the way they deliver them. Charleer et al. (2015) described the use of tabletop dashboards to provide an overview of students' access patterns in each learning activity and the time spent on each. Finally, Slotta et al., (2013) investigated the use of ambient analytics tools to support teachers to reflect on IED in physical learning environments, based (among others) on students' activity access patterns and time spent on each activity.

Overall, based on the above review of existing literature, it is argued that Task #2 is marginally supported by existing STEM-specific (or generic) TLA works. More specifically, the existing works aim to present students' activity in terms of individual educational design elements (for example, how much time did students spend on a specific resource). However, a different layer of insights, which is not yet accommodated, would be to extract latent characteristics of the IED, created by the combination of the individual elements. A prime example of such deeper analyses in the context of Inquiry-based STEM, would be to jointly analyze the educational tools used in an IED to elicit the level in which their combination provides adequate guidance to students for each Inquiry phase.

2.3 Research contribution

The research challenge that this work aims to address emerges as the holistic conceptualization of the challenges identified in the previous sections. In particular, it investigates whether Teaching Analytics can be used to assess IED (in terms of level and types of tool-supported guidance) and, also, whether such insights could provide a meaningful new lens to process and exploit data that are collected from existing Learning Analytics tools, ultimately informing systematic reflection on IED and provision of more tailored student experiences.

Therefore, the contribution of the paper is the design and thorough evaluation of a 'Teaching and Learning' analytics method (embedded within a supporting research prototype tool) that aims to accommodate both Tasks of STEM teachers. The proposed method offers a novel way to capture, assess and report the level and types of tool-supported guidance offered in IED **(Task #1)**. Also, it can receive students' educational data already collected from external tools, use that data to build student profiles and allow teachers to investigate the impact of different levels and types guidance on individual students and groups **(Task #2)**.

3. A Teaching and Learning Analytics Tool for Reflective Educational (re)Design

3.1 Overview

The proposed TLA tool (which hosts the proposed method) comprises two core components, the **Teaching Analytics** component and the **Reflection** component (Fig. 1):

The *Teaching Analytics* component (relevant to Task #1) refers to an analyzer which can be deployed on existing IED and assess them in terms of the tool-supported guidance they afford. The output of the *Teaching Analytics* component (defined explicitly below) is fed to the follow-up *Reflection* component.
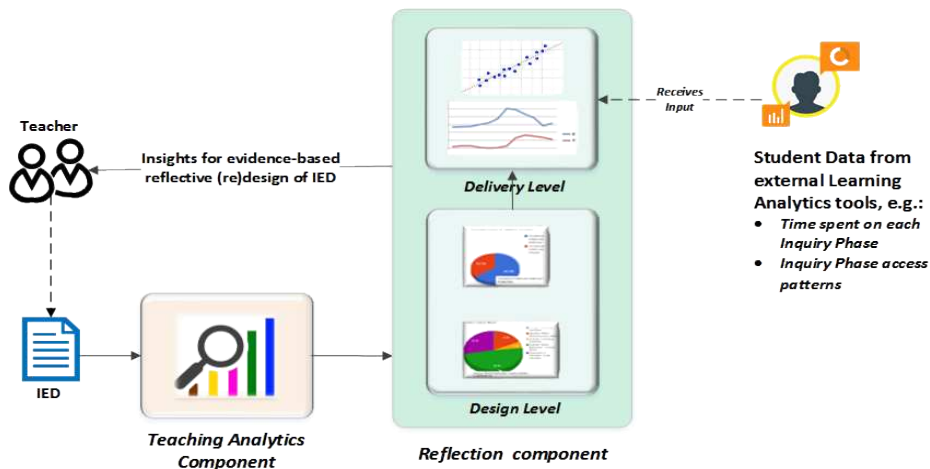
Figure 1: Overview of the proposed Teaching and Learning Analytics tool

- The *Reflection* component (relevant to Task #1 and Task #2) aims to support teachers to refine their IED. More specifically, the Reflection component provides dashboards depicting the guidance analysis results from the Teaching Analytics component at the *design level*, in various levels of granularity. Furthermore, the tool can also visualize correlation analyses between the guidance analysis results to customizable students' educational data from *the delivery level*, gathered from external learning analytics tools (e.g., embedded in digital learning environments). Finally, capitalizing on these data, the tool can also create and visualize student profiles in terms of guidance impact on their activity and/or performance.

Each of the two components of the proposed TLA tool are presented in detail in the following sections.

3.2 Teaching Analytics Component

The core engine of the proposed Teaching Analytics component is a mechanism used to evaluate the **degree of guidance** provided by the online labs and complementary guidance tools in an IED. The main rationale behind the algorithm is that online labs usually only partially afford the full range of Guidance Instances required in an Inquiry phase. Therefore, guidance tools are commonly employed in order to complement these design shortcomings.

In this context, the mechanism of the proposed Teaching Analytics component aims to calculate the level in which the selected guidance tools effectively *complement* the Guidance Instances already provided by the online lab of the IED. To perform this task, the mechanism scans the IED for inconsistencies between (a) the Guidance Instances provided by the guidance tools which have been selected by the teacher and (b) the Guidance Instances 'anticipated' in each Inquiry phase which are not accommodated by the online lab(s) used in the IED. The latter are elicited from training sets of exemplary IED with 'best practice' provision of guidance, which are used to configure the mechanism (described in Section 4.2).

Furthermore, each guidance tool is also assessed in terms of the level of match between the functionalities it affords (captured through keyword descriptions from the tool developers) and the textual description of the learning activities described within an Inquiry phase (captured as text-based outlines by the teacher).

Therefore, the proposed algorithm feeds on two types of input data. The first are the text-based outlines of the learning activities in each Inquiry phase of the IED (structured based on the Inquiry Cycle (Pedaste et al., 2015)). These outlines are provided by the teachers who design the IED. The second refers to the Guidance Instance profiles[4] and (keyword) descriptions of the online labs and guidance tools, utilized in the IED. The former profiles depict the Guidance Instances afforded by online labs/guidance tools and are generated offline. The latter descriptions are typically provided by the tool developers and present their core functionalities.

Capitalizing on these input data, the proposed Teaching Analytics component calculates the **Degree of Guidance** as a combination of two metrics:

- The *completeness* degree between the Guidance Instances offered by each guidance tool against the Guidance Instances required by the specific Inquiry phas*e*. In particular, the guidance tools are assessed in terms of the level that their Guidance Instance profiles complement the ones that the online lab(s) do not accommodate and, therefore, complete the Guidance requirements of the Inquiry phase. The completeness degree is calculated using the Jaccard similarity co-efficient, a well-known statistical method to accurately compare between attribute sets (e.g., Sergis & Sampson, 2016b).

- The *similarity degree* between the textual outline of each Inquiry phase (provided by the teachers) and the descriptions of the guidance tools (provided by the tool developers). To calculate the similarity degree, the *tf-idf* factor was used in combination with the cosine similarity factor (Lops et al., 2011). Capitalizing on these factors, the algorithm can elicit whether the specific functionalities of guidance tools match (and thus, support) the textual description of the learning activities of a given Inquiry phase.

The two aforementioned degree metrics are combined in order to generate the **Guidance Degree Index (GDI)** of the Inquiry phase (**IP**), depicting the level that the Guidance Instances required by the *IP* have been accommodated by the guidance tools (complementary to the utilized online lab).

---

[4] For this work, these profiles were created by the researchers following the detailed guidelines of Zacharia et al. (2015)
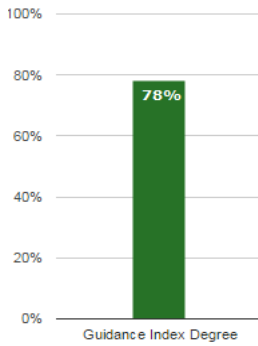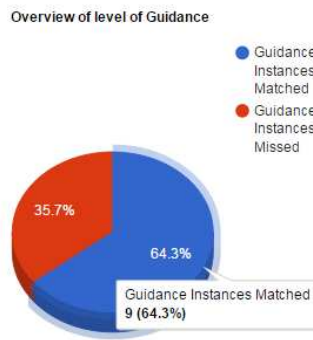
Figure 2(a): CGDI of IED

Figure 2(b): Cardinality of Guidance Instances accommodated and 'missed'
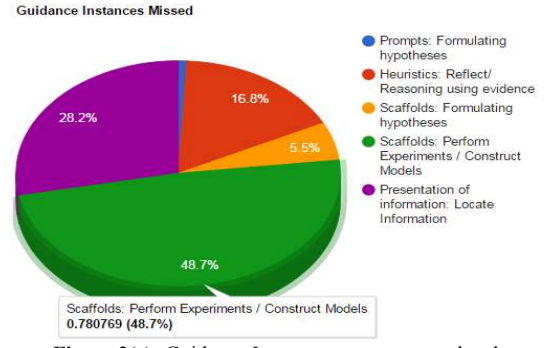

Figure 2(c): Guidance Instances not accommodated

The proposed mechanism calculates the **GDI** using the following conceptual formula:

$$GDI_{t,IP} = \lambda * \left(w_{GI} * Jaccard_{t,IP}\right) + (1 - \lambda) * \left(tf - idf_{t,IP}\right) \quad (1)$$

where:

- $Jaccard_{t,IP}$, is the Jaccard index calculated to represent the similarity between each Guidance Instance offered by the guidance tool t and the ones required by IP.
- w_{GI}, is a weighting factor that aims to consider the significance of each Guidance Instance G_I for the given IP. Such weighting factors need to be considered, based on the understanding that specific Guidance Instances might be more important to be included in a specific Inquiry phase than others. For example, providing tool-supported guidance for hypotheses formulation might be very significant to be provided during a hypothesis formulation task. The w_{GI} factor is elicited from the training set selected to train the proposed mechanism, which comprises exemplary IED with 'best practice' provision of guidance. The w_{GI} factor is calculated based on the past frequencies of each Guidance Instance for the corresponding Inquiry phase. Therefore, higher past frequency signifies higher importance of the corresponding Inquiry Instance for the specific phase.
- $tf - idf_{t,IP}$, is the degree of similarity between the keyword description of guidance tool t and the textual description of the IP based on the $tf\text{-}idf$ factor.
- λ, is a numerical parameter introduced in the function in order to calibrate the significance of each of the two factors contributing to $GDI_{t,IP}$. The value of this parameter (in terms of achieving highest accuracy) is obtained in an experimental manner and is based on the training set used to train mechanism. As it will be discussed in Section 4.1, obtaining this optimal configuration of the λ parameter is the explicit focus of the evaluation Layer 1.

Apart from calculating the GDI for each Inquiry phase of an IED, the Teaching Analytics component also generates a **Consolidated Guidance Degree Index (CGDI)** aimed at the overall IED. The CGDI is calculated as the mean of the **GDI**s of all Inquiry phases. Overall, the indicators that have been selected for calculating the (C)GDI are minimalistic, so as to ensure that
the results obtained are in line with the depth and reliability of the data that can be retrieved from different digital learning environments. Given that the granularity and richness of

educational data that can be collected from such environments is often heterogeneous and sometimes, incomplete (Kruger et al. 2010), ensuring access to rich data in all occasions cannot be ensured. Therefore, building on prior experience (Rodriguez-Triana et al., 2015), the proposed approach adopts a minimalistic
set of indicators that are, however, appropriate and sufficient to reveal facts that would otherwise be invisible for teachers.

In summary, at the design level, the proposed Teaching Analytics component incorporates a mechanism for analyzing existing IED in terms of the GDI and CGDI offered. These analyses are then fed to the Reflection component in order to support teachers' reflection, in two ways. First, via dashboards which are directly generated from the analysis results and assist teachers to gain visual overviews of their analyzed IED, at the design level.
Second, the outputs of the proposed Teaching Analytics component can provide the context to study diverse students' activity or performance indicators from the delivery of the IED (at the delivery level), as the latter can be extracted from existing Learning Analytics tools (e.g., from Learning Management Systems). This aims to provide teachers with the means for creating student profiles and investigating how different levels and types of guidance impacts the activity and performance of individual or groups of students. The *Reflection* component is presented in detail in the following section.

3.3 Reflection Component
The proposed Reflection component comprises a hub of functionalities that can provide teachers with feedback on their IED (with data from the design and delivery level) towards scaffold refinement of their IED. The functionalities offered can be divided into two levels, namely the design level and the delivery level.

3.3.1 Reflection Component: Design Level
At the design level, the proposed Reflection component aims to directly visualize the outputs of the Teaching Analytics component, in diverse levels of granularity. Different types of dashboards were explored in order to facilitate teachers to understand and gain insights on their IED (Verbert et al., 2014). The proposed Reflection component, at the design level, currently offers three types of dashboards, which are presented as follows:

- **Overall CGDI for the IED.** This dashboard aims to provide a summative overview of the Consolidated Guidance

Figure 3(a): Indicative correlation analyses for a student between the GDI for two Inquiry phases (Conceptualization and Investigation) and a student's time spent on these phases
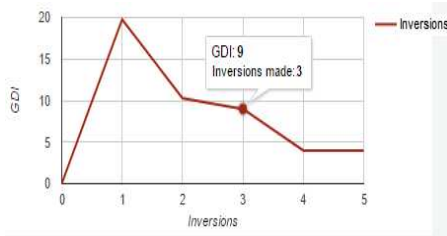
Figure 3(b): Indicative screen of aggregated students' (group) profile depicting connection between inversions and GDI of a specific Inquiry phase
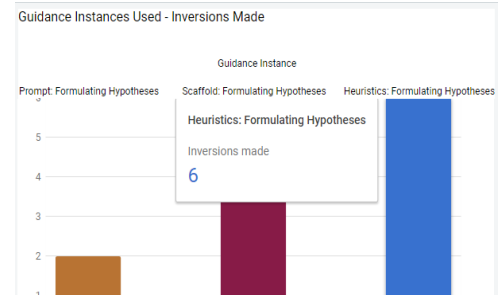
Figure 3(c): Indicative screen of student profile showing ranked connection between Guidance Instances used and inversions made for a specific Inquiry phase

Degree Index provided by the overall IED. Fig. 2(a) presents an exemplary depiction of this dashboard.

- **Specific Guidance Instances 'missed' in each Inquiry phase.** This dashboard provides a further layer of granularity, since it offers much more detailed information to teachers. More specifically, it provides a detailed depiction of the specific Guidance Instances that have not been incorporated in each Inquiry phase (but should have been based on the training set). Fig. 2(c) presents an exemplary deployment of the dashboards for one Inquiry phase. As the Fig. 2(c) depicts, each of the Guidance Instances which are not addressed in the given Inquiry phase are visualized in a variant degree of significance. This degree aims to represent the overall significance of the corresponding Guidance Instance for the given Inquiry phase. In this way, the Reflection component provides feedback on not only the specific Guidance Instances that they have not provided supporting tools for, but also on the importance of each one for the given Inquiry phase.

- **Cardinality of Guidance Instances accommodated and 'missed' in each Inquiry phase**. This dashboard provides an additional layer of granularity for the teacher. It aims to provide an overview of the cardinality of the Guidance Instance set that each Inquiry phase has been assigned to, as well as those that have not been incorporated (based on the insights elicited from the training set). Finally, the cardinality sets are also presented in percentage format. Fig. 2(b) presents an exemplary deployment of the dashboard for one Inquiry phase.

3.3.2 Reflection Component: Delivery Level

At the delivery level, the Reflection component aims to bridge the analyses of the Teaching Analytics component and relate them to students' educational data. In particular, in terms of educational data input, the research prototype tool can receive and process diverse types of data generated within digital learning environments (such as Learning Management Systems) using .json files. The selection of the data to be provided to the tool depends on the teachers' focus of inquiry as well as the availability of student activity and performance data that the digital learning environment harvests. Examples of student data that the TLA tool can exploit include students' access patterns in the learning activities of an IED or the time spent on each Inquiry phase. The Reflection component adopts this 'generic' approach in terms of educational data, to accommodate more degrees of freedom for teachers. More specifically, teachers' capacity to reflect on their practice can be restricted when only one source of student data is available (Dyckhoff et al., 2012;

2013). Therefore, the Reflection component aims to provide them with more freedom, by allowing them to utilize student data relevant to their own teacher inquiries.

Building on the student data and the IEDs stored within the digital learning environment, the tool generates a set of insights for the teacher. In particular, the tool can calculate the correlations between the student data and the (C)GDI of the IED that students were engaging with. These analyses can facilitate teachers to investigate whether and how the level of guidance (and specific Guidance Instances) is related to students' activity and performance. The tool can visualize these analyses to the teachers using dashboards, for both individual students as well as for groups of students. Fig. 3a presents an indicative dashboard depicting the results of correlation analyses for a student between the GDI of two exemplary Inquiry phases (Conceptualization and Investigation) and the students' time spent on these phases. As aforementioned, different types of student data and analyses regarding different phases can be used to generate these dashboards.

Furthermore, the TLA tool can also process the student activity (and/or performance) data in order to formulate **profiles** of individual or groups of students, in terms of guidance impact on activity and/or performance. More specifically, these profiles comprise duplets of students' data in each phase of an IED (and overall IED) and the (C)GDI of this phase (or full IED). Group profiles are similar, but in terms of student data, they contain the average value (where relevant) for the students of the group. For example, Fig. 3b depicts an indicative dashboard of aggregated students' profile data showing the connection between the GDI of a specific Inquiry phase and the number of inversions that these students made for this Inquiry Phase (i.e., when a student did not follow the ordering of Inquiry phases that the teacher had defined).

Additionally, the profiles can also capture duplets of the Guidance Instances in each phase with the students' data for this phase. Capitalizing on these duplets, the TLA tool can generate dashboards that visualize the impact patterns of different levels and instances of guidance on students' activity, for example which level (or instance) of guidance was associated with a student's highest performance. In this way, teachers can identify which levels or types of guidance appear to be most beneficial for individual or groups of students. For example, Fig. 3c depicts an indicative screen of student profile showing the inversions made by a student for an Inquiry phase (across different IED) mapped to the Guidance Instances used (ranked ascending since fewer inversions suggest a smoother learning experience). The dashboard shows that the most inversions were made when the student was supported with a

Heuristics type of guidance when formulating their Inquiry hypotheses.

All insights offered by the Reflection component at the delivery level aim to help teachers infer how the provided level and instances of guidance can impact individual and groups of students' activity and/or performance. The teacher can exploit these insights and refine their IED design, so as to accommodate more appropriate level and types of Guidance for their students.

It is noted that even though the TLA tool is planned to be adjusted to receive .json data files from diverse digital learning environments as part of future work, for the purpose of this work the research prototype tool has been configured for a specific case study, in order to generate a robust proof-of-concept through an extensive evaluation protocol. This evaluation protocol, which is described in the following section, was conducted within the context of the Go-Lab project. Therefore, the research prototype tool has been configured to receive IED and students' activity data from the "Graasp" digital learning environment (Vozniuk et al., 2016), which was adopted by the project.

3.3.3 An indicative scenario of use

This section provides a brief conceptual scenario of use of the proposed TLA method and tool. The workflow starts with the teacher designing their IED in the digital environment. In the case study adopted in this paper, this environment is the Graasp system and the proposed research prototype tool has been plugged-in to this system. The Graasp system offers teachers with a wizard-like process for designing IED following a template implementing the Inquiry Cycle of reported in Pedaste et al. (2015). During the design of the IED, teachers can also populate each phase with resources, as well as guidance tools and online labs, available from the Go-lab repository [http://www.golabz.eu] (in this case study).

After the IED has been designed and populated, the Teaching Analytics component of the proposed tool (which is connected to the Graasp system database) can be deployed to analyze the IED. The tool retrieves the full IED (which is formatted and shared as a .json file) and also retrieves the metadata of the guidance tools and online labs used in the IED (also through .json files). Building on this information, it employs the process presented in section 3.2, and calculates the (C)GDI for this IED.

Following this part, the Reflection Component (at design level) is automatically triggered so as to present teachers with the outcomes of the analysis using the dashboards outlined in section 3.3.1. In this way, the teacher receives a detailed overview of their IED in terms of provided guidance and may choose to edit some of the design choices they made, for example, replace specific guidance tools to increase the GDI of a particular Inquiry phase before they deliver it to their students.

After finalizing the IED, the teachers can deliver it to their students, through a learning management system (in this case, utilizing the delivery functionalities of the Graasp system). The Graasp system collects different types of educational data from students during the delivery of IED (see section 4.1.2 for indicative examples). These data are stored and can be accessed by the proposed tool. Therefore, after the end of the IED

delivery, the teacher can re-deploy the tool and trigger the Reflection component (at delivery level). In particular, they can select different types of data that are already collected for individual students (such as the time spent on each inquiry phase and tools or the sequence of phases that was followed). The proposed tool automatically retrieves the selected data and also retrieves the analysis results for the IED that was delivered (Teaching Analytics component). By jointly processing these two sources of data, the tool generates the dashboards presented in Section 3.3.2. For example, it investigates correlations between the GDI of phases to students' frequency of inversions; or builds individual student profiles regarding frequency of inversion or time spent on different phase. By having access to such insights, teachers can understand how different students respond to variant levels and types of guidance. This information could lead to them designing more personalized experiences, by providing individual (or groups of) students with the most appropriate level of guidance to optimize their performance and engagement.

The following section outlines the evaluation methodology adopted in this work.

4. Evaluation

4.1 Evaluation Methodology

A two-layer evaluation methodology was adopted in this work. Each layer of the methodology aims to evaluate different aspects of the TLA tool, as follows:

- **Layer 1** aimed at evaluating the accuracy of the TLA tool mechanism to identify guidance shortcomings in IED at the design level (*Teaching Analytics* component – relevant to Task #1).
- **Layer 2** focused on evaluating whether the (C)GDI (as generated by the TLA tool) offers a statistically significant indicator related to students' activity during the delivery of the IED (*Reflection* component at the delivery level – relevant to Task #2).

Each of the two evaluation layers is presented in detail in the following sections.

4.1.1 Layer 1 Evaluation Methodology

The methodology employed for the Layer 1 evaluation focused on the accuracy of the proposed TLA to calculate whether the selected guidance tools effectively complemented the guidance needs of a given IED and its phases.

Furthermore, the methodology also aimed at configuring the mechanism of Formula (1) in terms of the value of parameter λ to achieve highest accuracy. This configuration was then exploited in the follow-up Layer 2 evaluation.

An instance of the Layer 1 evaluation has already been performed and reported in (Sergis & Sampson, 2016a). In the context of this work, however, this initial evaluation was extended with additional novel metrics to further validate the results, in terms of precision and overall accuracy. The following set of widely-used metrics was exploited (elaborated in Sokolova & Lapalme, 2009):

- **[M1] - Precision.** This metric is defined as the ratio of the correctly classified "positive" elements (in this case: 'appropriate' guidance tools) by the mechanism against all the "positive" elements classified by the mechanism (both

correctly and falsely).

- **[M2] - Recall.** This metric is defined as the ratio of the actual "positive" elements classified by the mechanism against all "positive" elements available in the full pool of elements. This metric was included in the initial instance of the Layer 1 evaluation.
- **[M3] - F1 measure.** This metric is aimed to combine the Precision and Recall metrics, towards offering a combinatory means to evaluate the mechanism.
- **[M4] - Specificity**. This metric is defined as the ratio of the actual "negative" elements classified by the mechanism (in this case: 'non-appropriate' tools) against all "negative" elements available in the full pool of elements. This metric was included in the initial instance of the Layer 1 evaluation.
- **[M5] - AUC (Area Under the Curve).** This metric is aimed at evaluating the sensitivity of the mechanism to avoid wrong classification of elements. This metric was included in the initial instance of the Layer 1 evaluation.
- **[M6] - Accuracy.** This metric aims to calculate an aggregated measure of the effectiveness of the mechanism, defined as the ratio of the correct classifications of the mechanism (both "positive" and "negative") against all classifications made.

4.1.2 Layer 2 Evaluation Methodology

Evaluation Layer 2 aimed to investigate whether the outputs generated by the TLA method (namely, (C)GDI) provide statistically significant indicators that can impact students' activity in the delivery of IED.

In terms of student data to be included in the evaluation, we were interested in studying *student activity* during the delivery of IED, therefore students' performance data (e.g., assessment scores) were not accommodated. Future evaluation studies will specifically incorporate students' performance data, as well. Therefore, capitalizing on the relevant literature, a set of data types was adopted, as follows:

- *Students' access patterns* in Inquiry phases of the IED. These data were utilized since existing literature suggests that the provision of tool-supported guidance to students could facilitate them to follow the initial ordering of the IED and avoid overriding it (namely, perform **inversions**), leading to a smoother learning experience (Manske et al., 2015).
- *Students' time spent on each Inquiry phase* and the overall IED. These data were utilized as a proxy of student perseverance, since literature suggests that the provision of tool-supported guidance can facilitate students to persist when engaging in Inquiry tasks and, consequently complete them (e.g., Girault & Ham, 2014).

Capitalizing on these data, Layer 2 evaluation focused on studying students' behavioral patterns and access distributions in relation to the generated GDI and CGDI, consistent with the evaluation protocols and analyses adopted in the literature works discussed in section 2.2. In particular, Layer 2 comprised four strands of investigation:

1. Investigate correlations between the [overall IED] and [per Inquiry Phase] Guidance Degree Index generated by the TLA tool and the students' runtime inversions in the corresponding IED and Inquiry Phases.

2. Investigate correlations between the [overall IED] and [per Inquiry Phase] Guidance Degree Index generated by the TLA tool and the students' time spent on the corresponding IED and Inquiry Phases.
3. Investigate the impact of the Inquiry phase Guidance Degree Index (as generated by the proposed TLA tool) on the distributions of students' access behaviors (i.e., the Inquiry phases they accessed).
4. Investigate the impact of the Inquiry phase Guidance Degree Index (as generated by the proposed TLA tool) on the students' access patterns (i.e., the flow of access in Inquiry phases).

4.2 Datasets

In order to perform the Layer 1 and Layer 2 evaluations, data from the Go-Lab project were used. More specifically, as it will be described in detail below, the data comprised 37 IEDs from instructional design experts and also 98 IEDs from school teachers designed in the context of the project as well as educational data from the delivery of these IEDs to 1832 students in various European schools. Furthermore, the Go-Lab online repository [http://www.golabz.eu] boasts a vast pool of 403 online labs and 41 guidance tools[5], which were exploited in the design and deliver of these IED. All aforementioned data used were collected during the overarching project evaluation activities (i.e., not specifically for this evaluation methodology), therefore they provided an objective means to evaluate the proposed method. A detailed description of the IED and educational data collected from the Graasp system can be found in Rodríguez-Triana et al. (2014).

Table 2 provides an overview of the evaluation methodology described in this section.

Table 2: Overview of Evaluation Methodology

	Evaluation Methodology		
	Benchmark	Focal Point	Metrics
Evaluation Layer 1	37 Expert-created, best-practice IED	Accuracy of the mechanism to identify (C)GDI design shortcomings	Precision, Recall, F1, Specificity, AUC, Accuracy
Evaluation Layer 2	98 Teacher-created IED	(C)GDI (as output of tool) impact on students' activity	Correlation Analysis, Chi-square test, lag Sequential Analysis

Regarding the **Layer 1 evaluation**, as discussed in Sergis & Sampson (2016), the training set comprised a Guidance Instance model derived from a state-of-the-art literature analysis (Zacharia et al., 2015), presented an outline of how Guidance Instances have been utilized in technology-supported IED and also provided quantitative analyses of their findings corresponding to each Inquiry phase (also consistent with the work of Pedaste et al., 2015). The test set was harvested from the publicly available repository of Go-Lab, containing 37 best-practice IED created by expert project partners in the field of Inquiry-based STEM education. These best-practice IED (which were populated with online labs and complementary

[5] Data from November 2016

guidance tools) were used to evaluate the accuracy of the proposed Teaching Analytics mechanism to correctly classify guidance tools as appropriate for given IED and its phases. In particular, first, the utilized guidance tools in the best-practice IED were 'hidden'. Then, the Teaching Analytics mechanism was deployed on the full pool of guidance tools available in the Go-Lab online repository, evaluating their capacity to complement the affordances of the online lab of the IED and meet the guidance needs of each IED phase. The top-ranked selections of the mechanism were then compared to the actual tools used in the IED (which were initially 'hidden'). Finally, the performance of the mechanism was evaluated based on these comparisons, using the diverse metrics discussed in section 4.1.1.

Regarding **Layer 2 evaluation**, the *training set* consisted of the 37 best-practice IED used in Layer 1. Additionally, the proposed TLA mechanism adopted the optimal configuration of Formula (1), as this was derived from Layer 1. The *test set* used comprised 98 teacher-created IED, utilizing online labs and guidance tools offered in the Go-lab repository, delivered to 1832 students. During these deliveries, which were part of the project evaluation activities, educational data from students were collected using an Inquiry-based teaching and learning environment used in the project (Graasp) (Gillet et al., In Press), including the data utilized in the evaluation protocol, namely (a) students' access patters in Inquiry phases of the IEDs and (b) students' time spent on each Inquiry phase and the overall IEDs.

5. Results

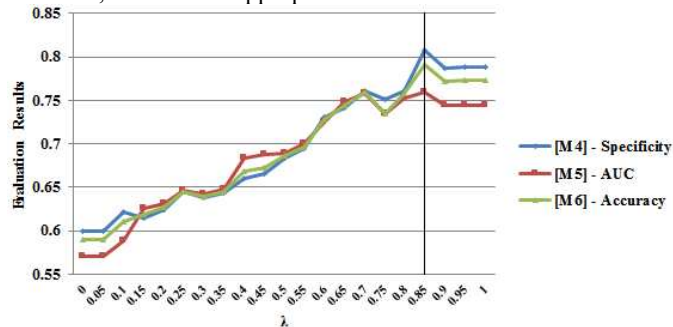## 5.1 Evaluation Results: Layer 1

The results of the Layer 1 evaluation based on the six metrics are presented in Fig. 4 and Fig 5. Fig. 4 presents the results for the evaluation metrics M1-M3. As the Fig. 4 depicts, the proposed mechanism obtains its highest accuracy value of Formula (1) for all three evaluation metrics for λ=0.85. More specifically, the results for the Recall metric suggest a high level of sensitivity for identifying the guidance tools that are appropriate for a given phase. This means that the mechanism is accurate in identifying the guidance tools (from the overall pool of available tools) whose Guidance profiles and functionalities are appropriate for supporting the Guidance needs of a given Inquiry phase.

Moreover, the results of the Precision metric highlight that the mechanism is significantly accurate in identifying which guidance tools are appropriate for a given Inquiry phase, considering as 'positive' elements only the guidance tools which were included in the best-practice IED (i.e., and not other guidance tools in the overall available pool which could have also been potentially appropriate). Finally, the consolidated results of the F1 measure, which provides a combinatory metric for Recall and Precision, further supports the high levels of sensitivity of the proposed mechanism, especially for the optimal point of λ=0,85.
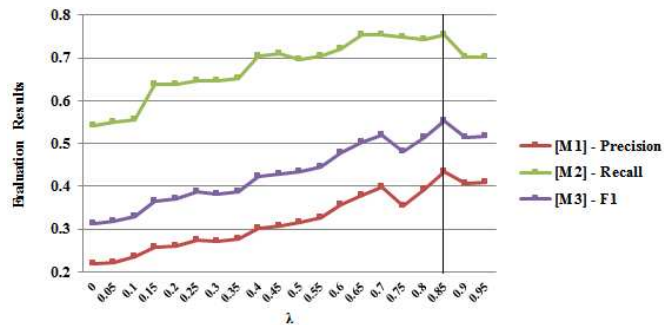
Figure 4: Evaluation results for the proposed TA tool in terms of Precision, Recall and F1 Measure.

Fig. 5 presents the results for the evaluation metrics M4-M6. As the Fig. 5 depicts, for all evaluation metrics M4-M6, the proposed classifying mechanism obtains its 'optimal' values for λ=0.85, similarly to the evaluation metrics M1-M3. Moreover, the evaluation results for the Specificity metric show that the proposed mechanism is very accurate at classifying guidance tools based on the capacity of their Guidance Instance affordances and functionalities to address the specific guidance requirements and outline of the given Inquiry phases and therefore, outline non-appropriate tools.

Figure 5: Evaluation results for the proposed TA tool in terms of Specificity, Area Under Curve and Overall Accuracy Measure.

The results of the AUC metric show that the proposed mechanism can effectively assess the capacity of guidance tools to better complement the Guidance Instances provided by the online lab(s) selected and, therefore, can avoid erroneous classifications at a very high degree.

Finally, the overall Accuracy metric supports the insights gained from all the previous five metrics. Therefore, it further fortifies the standpoint that the proposed mechanism can effectively assess the appropriateness and completeness of the Guidance Instances provided through guidance tools for a given IED and its Inquiry phases.

Capitalizing on the promising findings from Layer 1 evaluation, which defined the optimal configuration of the proposed classifying mechanism (i.e., λ=0,85) and attested to the overall high levels of accuracy of the mechanism of the proposed Teaching Analytics component, the Layer 2 of the evaluation was performed.

5.2 Evaluation Results: Layer 2

The Layer 2 of the evaluation aimed at investigating the relation between the outputs of the tool (namely the GDI and CGDI) with student activity data from the delivery of the IEDs, in order to assess the tool's capacity in supporting teachers' (re)design. The following sections present the results obtained for each of the four strands of investigation defined for this

evaluation layer.

5.2.1 Evaluation Results: Investigation strand 1

Investigation strand 1 aimed to investigate, using correlation analysis, the potential link between the (C)GDI and the students' runtime inversions in the corresponding IED and Inquiry phases. These inversions refer to whether the students accessed the IED in a different sequence than the initial ordering of the IED phases as defined by the teacher. Table 3 presents the results obtained. It is noted that the Inquiry phases follow the typology outlined in Pedaste et al. (2015).

As the Table 3 depicts, there is a weak, but statistically significant, negative correlation between the overall IED Guidance Degree Index generated by the TLA tool and students' percentage of inversions made in the corresponding IED. This indicates that the highest the Guidance Degree Index identified, the least the percentage of runtime inversions demonstrated by the students during the delivery of the IED.

Table 3: Pearson's Correlation Between Calculated Guidance Degree Index and Students' Inversion Percentage

	Pearson's correlation	p
Orientation [O]	-0.35**	0.01
Conceptualization [C1]	-0.41*	0.03
Investigation [I]	-0.38**	0.01
Conclusion [C2]	-0.32*	0.03
Discussion [D]	-0.16	0.15
Overall (per IED)	-0.20*	0.04

*p<0.05, **p<0.01

Furthermore, in terms of specific Inquiry phases, similar negative correlations were identified. More specifically, a weak but statistically significant correlation was identified for the Conclusion Phase. Additionally, moderate and statistically significant correlations were highlighted for the Conceptualization, Investigation and Conclusion Phases. This means that for the latter three Inquiry phases, providing a high level of Guidance was a significant factor for assisting students to follow the designed flow of phases and not engage in unexpected deviations and/or inversions. Finally, in terms of the Discussion Phase, no statistically significant correlation was identified. This can be explained by the fact that in the Discussion phase, students are usually expected to reflect on and discuss their experimental results and process. Therefore, visiting previous phases can be regarded as an intrinsic aspect of this phase, regardless of the Guidance Degree Index provided.

Overall, the findings for Investigation strand 1 provide evidence on a clear connection between the outputs of the TLA tool and students' access patterns (in terms of inversions). Therefore, based on these insights, it is reasonable to argue that teachers could exploit the insights generated to refine their IED and achieve more coherent student engagement in terms of flow of Inquiry phase access.

5.2.2 Evaluation Results: Investigation strand 2

Investigation strand 2 aimed to study potential correlations between the (C)GDI generated by the TLA tool and the students' time spent in the corresponding IED and Inquiry phases. Table 4 presents the results obtained.

Table 4: Pearson's Correlation Between Calculated Guidance Degree Index and Students' Time Spent

	Pearson's Correlation	p
Orientation [O]	0.16*	0.05
Conceptualization [C1]	0.35**	0.00
Investigation [I]	0.39**	0.00
Conclusion [C2]	0.10	0.22
Discussion [D]	0.21*	0.03
Overall (per IED)	0.21**	0.00

*p<0.05, **p<0.01

As the Table 4 depicts, there is a weak, but statistically significant positive correlation between the overall IED Guidance Degree Index and the students' overall time spent in the corresponding IED. This suggests that the highest the GDI, the more time students actually commit to engaging with it. This is consistent with findings from other works (e.g., Girault & Ham, 2014) where higher levels of guidance provision resulted in increased time engagement and, subsequently, to better student performance.

Furthermore, in terms of specific Inquiry phases, statistically significant positive correlations were identified for the Orientation and Discussion phases (weak correlations) and the Conceptualization and Investigation phases (moderate correlations). This finding can be explained by findings from other works, which argue in favor of the positive impact of tool-supported guidance on students' capacity to engage in tasks commonly addressed in these phases, for example (Kim & Pedersen, 2011). Therefore, higher levels of guidance tend to facilitate students to persist when engaging with such tasks. Finally, in terms of the Conclusion phase, no statistically significant correlation was identified. This finding can be explained by the conclusions from the upcoming investigation strands 3 and 4, where it is evident that the Conclusion phase was usually either (a) not systematically accessed by the students or (b) was accessed in a manner not consistent with the initially designed ordering of the IED, therefore students potentially did not fully engage in the relevant tasks.

Overall, the findings argue that the outputs of the TLA tool are significantly related to the students' time commitment to engaging with the IED and its phases. Similar to Investigation strand 1, the results suggest that the tool can provide meaningful insights to teachers in order to provide appropriate guidance to the students that will facilitate them to persist on engaging in the IED phases.

5.2.3 Evaluation Results: Investigation strand 3

Investigation strand 3 aimed to investigate the relation between the Inquiry phase (C)GDI generated by the TLA tool and the distributions of students' level of access in each Inquiry phase, i.e., 'access behaviors' of which Inquiry phases were accessed by the students independently from the sequence in which this was performed). In order to address this strand, the test set of 98 teacher-created IED was divided in two clusters based on their Inquiry phases' GDI, utilizing the widely-used *k-means clustering algorithm*. Based on the results of the clustering process, Cluster 0 contained 50 IED attributed with 'Low' GDI and Cluster 1 contained 48 IED attributed with 'High' GDI.

Following this clustering, the students' access behavior in each of the two IED clusters were identified. To do that, the distribution and frequency of their access behaviors in terms of Inquiry phases was harvested and processed, following the methodology described in Yang et al. (2015). The harvested data were normalized for each cluster in terms of the student cardinality in each cluster. The final, normalized results of this process are shown in Table 5.

Table 5: Normalized Frequencies of the students' access behavior in each Inquiry Phase

	Students	O	C1	I	C2	D
Cluster 0	848	28,93	8,38	6,77	3,47	3,07
Cluster 1	984	39,20	9,57	11,37	4,96	3,94

A chi-square test was conducted on these data to identify potential significant differences in the behavior distributions between the students in each of the two clusters. The result ($\chi^2(4) = 6,02, p < 0.00$) indicated that the student behavior distributions in terms of Inquiry phase access were different in a statistically significant manner.

More specifically, the main finding is that the students engaged with IEDs from Cluster 1 were universally presenting higher levels of access in all Inquiry phases than those engaged in IEDs of Cluster 0. This means that students engaged in IEDs from Cluster 0 presented a larger tendency to skip and omit phases, especially the Conclusion and Discussion phases. On the other hand, IEDs from Cluster 1 were more engaging to students in the sense that they supported them to complete the full range of phases designed by the teacher.

Overall, the data suggest that the outputs of the proposed TLA tool are statistically related to the students' access behaviors in Inquiry phases in a significant manner. Therefore, the results for Investigation strand 3 show that the proposed TLA tool provides insights that are meaningful to facilitate students in engaging in and completing the full length of the IED. Moreover, capitalizing on these findings, it is reasonable to argue that teachers can use the proposed TLA tool in their teacher inquiries, by exploiting the student profiling capabilities it affords. In particular, these findings indicate that capitalizing on the student profiles, a teacher could provide more tailored guidance to individual (or groups of) students, leading to more effective learning experiences. However, this standpoint should be explicitly investigated in future works, as discussed in the concluding section.

5.2.4 Evaluation Results: Investigation strand 4

Investigation strand 4 aimed to investigate the relation between the (C)GDI and the students' flow of access in Inquiry phases, namely the path that students follow within an IED. For this strand, the previously described clustering of the IED was again exploited. To answer this strand, lag Sequential Analysis was employed to elicit statistically significant student access behaviors when engaging with the different phases of the IED (Yang et al., 2015). More specifically, lag Sequential Analysis was utilized to study the flow in which students accessed Inquiry phases and draw conclusions about how these flows are related to the GDI. This method was selected since it has been previously successfully exploited to investigate the different sequences of activities between students (e.g., Hou, 2015).

Following the methodology of lag Sequential Analysis, (e.g.,

Yang et al., 2015), the frequency transition tables were calculated, namely the frequency of students' access behavior in each Phase immediately following another phase. Based on this, the resulting adjusted residuals tables were calculated based on the z-score (Table 6 and Table 7) (Wang et al., 2014), so as to examine the students' behavioral patterns of accessing different Inquiry phases and identify potential statistically significant differences in each cluster. A z-score value of each sequence greater than +1.96 (depicted in green highlight in Table 6 and Table 7) indicated that a sequence reaches the level of significance (p < 0.05), as defined by Wang et al., (2014). Fig. 6 and Fig. 7 depict the access patterns that have reached statistical significance based on the z-score of each sequence. The arrows indicate the direction of each sequence, while the associated numbers describe the strength of the transition probability.

Table 6: Students' access patterns for each Inquiry Phase (adjusted residuals for Cluster 0)

	O	C1	I	C2	D
O	0,356	*15,029*	*2,801*	-9,698	-8,489
C1	-6,366	*9,697*	-7,626	*12,099*	-7,804
I	-7,682	-5,822	*12,198*	*9,526*	-8,219
C2	-8,686	-9,357	*2,050*	0,754	*15,239*
D	-5,981	-6,669	-6,446	*2,625*	*16,471*

Overall, the data related to IEDs from Cluster 0 (Fig. 6) present students' access patterns as a fragmented flow. This means that students tended to override the initially designed flow of phases in IEDs of Cluster 0. This is evident from (a) the lack of statistical significant transitions between consequent phases (i.e., C1-> I) and (b) the statistically significant inversions which are identified, namely O->I, C2-> I and D->C2. The former means that students tended to skip the Investigation phase when originating from the Conceptualization phase. The Investigation phase was instead mostly visited either from the Orientation phase (thus, skipping the Conceptualization phase) or from the Conclusion phase, namely, students navigated the IED in a manner which was not aligned to the ordering of the teacher design.
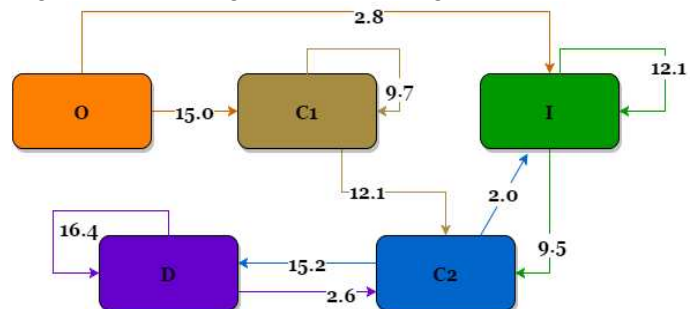
Figure 6: Students' access patterns for IED in Cluster 0 (Low GDI).

Furthermore, there appear to be statistically significant student patterns 'repeating' specific phases (i.e., C1, I, D). Such behaviors can be potentially attributed to students' reviewing their activities and work within a phase, and should not be considered as divergent behavior from the initially planned designs. Overall, the analysis of data from the Cluster 0 IED provides evidence that students only partially followed the

teachers' designs and often engaged in both inversions as well as omissions of specific phases.

On the contrary, data from IEDs belonging to Cluster 1 (Fig. 7) signify a more structured and organized flow of access patterns, in terms of the initially designed IED.

Table 7: Students' access patterns for each Inquiry Phase (adjusted residuals for Cluster 1)

	O	C1	I	C2	D
O	*6,822*	*13,631*	-2,511	-9,407	-8,536
C1	-6,704	*7,645*	*13,741*	-7,315	-7,367
I	-7,246	-5,698	*15,122*	*5,391*	-7,568
C2	-6,662	-6,572	-4,031	0,055	*17,209*
D	-4,871	-5,336	-5,008	-2,565	*17,780*

More specifically, the data delineate a statistically significant path that transverses the full IED in the initial ordering of the Inquiry Cycle. This means that students engaged with the IEDs in a manner compliant to the teachers' design, and avoided omitting phases or visiting phases in an ad-hoc manner. Moreover, a pattern of specific phase repetitions was also identified (i.e., O, C1, I, D), similar to the findings from Cluster 0. As aforementioned, these behaviors (which do not alter the designed ordering of the Inquiry phases) are likely to be related to students' reviewing a specific phase.

Overall, the results for investigation strand 4 signify that the proposed TLA tool can provide teachers with meaningful feedback towards reviewing and altering their IED, since the outputs of the tool were found to be significantly influential on students' access pattern coherence.
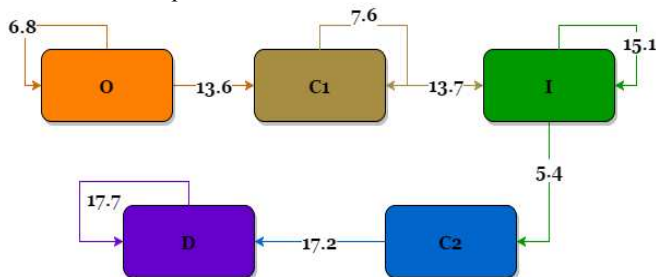
Figure 7: Students' access patterns for IED in Cluster 1 (High GDI)

Furthermore, similar to strand 3, these findings argue for the potential of the proposed TLA tool to support teachers build more tailored learning experiences for their students, capitalizing on the student profiles built. By measuring (and visualizing) how individual (or groups of) students' learning paths are affected by the provided guidance, teachers could better meet their students' particular guidance needs. However, it is again noted that these hypotheses would need to be explicitly investigated and validated in future work.

6. Discussion

The rationale for this work stems from the need to facilitate teachers in the process of reflection and (re)design of their IED. To meet this need, the core standpoint is that educational Analytics methods should move beyond tracking and reporting student activity and performance for the purpose of individual student scaffolding during the learning process (primary focus of existing Learning Analytics approaches). Instead, these methods need to support teachers with insights on how to *systematically* improve their teaching practice and provide more tailored and effective learning experiences to the students. Such

insights should support teachers in analyzing as well as evaluating their educational designs, based on data-driven insights.

The envisaged added value of the proposed method (and tool) primarily derives from the different levels of applicability, both in terms of informing other research works related to educational data Analytics as well as improving teachers' practice.

In terms of research, the proposed approach and the insights from this work in general could be used to enhance relevant strands aiming to investigate methods and tools to strengthen teachers in analyzing and improving their teaching practice with educational data. For example, Rienties & Toetenel (2016) have been studying ways to analyze educational designs in Higher Education contexts, so as to explore how different types of learning activities correlate with students' performance. Similar studies are being conducted by Nguyen et al. (2017). The method reported in this paper (despite currently being IED-specific) could enhance this research strand, by offering another level of analysis of educational designs focusing on the affordances of the resources and tools included. This level of analysis would supplement that of learning activities and potentially lead to more granulated insights when correlated with students' performance data (or other data, such as engagement, satisfaction and grit).

Another research strand that could benefit from this work refers to works that have been striving to use diverse educational data in order to help teachers track and improve their own performance when delivering their practice (i.e., moving beyond reflection on design, to reflection on delivery). A core aspect in these works is a detailed analysis of the educational design, in order to use educational data from the delivery for self-'monitoring' the performance of teachers and support their reflection (e.g., Monroy et al., 2014; van Leeuwen et al. 2014; Rodríguez-Triana et al., 2015). Therefore, the ideas behind the proposed method in this paper are consistent with this strand of research and could introduce new perspectives of self-'monitoring', for example, on how to optimize the level and type of guidance that teachers provide to students on-the-fly.

Finally, another research strand that could benefit from the proposed approach refers to the works investigating how to effectively provide guidance to students in online environments. Despite the existing body of research revealing the potential of computer-supported guidance in on-line contexts to promote students' self-regulation and problem-solving (Belland et al., 2017), the challenge of effectively eliciting students' guidance needs and offering personalized guidance in such context has been scarcely studied. This can prove a significant shortcoming, considering the large and diverse cohorts of online students which can make it difficult for teachers and tutors to identify and meet their individual guidance needs. Therefore, the proposed method (and tool) could provide a useful means to progress such research agendas, by allowing a transparent view to how each student (or cohort) utilizes and is affected by the guidance provided in the educational design.

In terms of informing teacher practice, the proposed method and the final 'end-user oriented' version of the tool (including front-end interfaces; to be developed as future work) could provide a means to facilitate teachers streamline the processes

of analyzing and interpreting diverse types of educational data; including data from their IED and their students (at individual and group level). As it is discussed further, longitudinal studies have been planned with teachers so as to evaluate the impact of the proposed method and tool on their day-to-day practice. It is envisaged that capitalizing on the lessons learned from these studies, the final version of the method and tool will offer a meaningful way to unravel previously hidden, or difficult to retrieve, aspects of both the teaching and learning process. In that way, teachers will have the capacity to access automatically-generated analyses and insights on their IED, and building on that, provide their students with more tailored learning experiences based on individual guidance profiles. Finally, as the educational paradigm continuously shifts to online environments (even in K-12 education), the proposed method (which is primarily capitalizes on educational data from digital learning environments, such as Learning Management Systems) could also assist teachers to better deliver and reflect on their IED in such environments; a task that can be considered cumbersome since there is limited direct interaction with students in order to detect potential challenges they face and decide on appropriate guidance to offer.

7. Conclusions and Future Work

This paper presented the design of a novel TLA method and research prototype tool, which aim to address a widely-acknowledged challenge in the context of STEM Education. More specifically, the proposed TLA tool aims to allow teachers to analyze their existing IED in terms of tool-supported guidance, which is a significant factor for effective IED delivery and relate these analyses to customizable students' educational data (which are collected externally) to facilitate the re-design process.

We introduced a two-layer evaluation protocol, aiming to assess the capacity of the proposed TLA method/tool to analyze existing IED and, also, to robustly validate whether and how the provided insights of the tool offer significant indicators influencing students' activity. The results showcase that the proposed TLA method and tool generate insights that are significantly correlated with diverse aspects of students' activity during the Inquiry learning process. Therefore, it is argued that the proposed TLA tool could be used to support teacher inquiries aiming to improve students' learning experiences through the provision of appropriate and, potentially, personalized guidance.

Future work in this agenda is envisaged to span different dimensions. First, it should aim to explicitly address the main limitation of this study, which is the framing of the evaluation within the context of Go-Lab, adopting the conventions regarding student and IED data format and descriptions' availability. Follow-up work should aim to explore means to generalize the potential of the research prototype tool to collect and process educational data through more generic approaches, such as linked data and social semantics (e.g., Ruiz-Calleja et al., 2012).

Additionally, deeper longitudinal evaluation studies need to be conducted through teachers' inquiries in order to provide robust evidence on the added value of the TLA tool in terms of improving teachers' reflection and enhancing students' engagement and performance in Inquiry-based STEM education. These studies need to explicitly evaluate the potential of the "Reflection" component of the proposed tool, by adopting both quantitative and qualitative methods. Additionally, these studies should also focus on collecting diverse quantitative and qualitative educational data regarding various students' performance dimensions, such as cognitive learning outcomes, Inquiry and problem-solving skills as well as affective (such as stress or frustration) characteristics. The potential and impact of the student 'Guidance profiles' offered by the tool need also to be investigated as a means to support teachers' offer more effective learning experiences for individual students or groups.

Furthermore, more longer-term future work can focus on merging this research strand with the emerging field of Affective Computing (D' Mello & Kory, 2015). This synergy of methods and tools could provide teachers with more detailed insights on how guidance affects students' emotional and self-regulative states during the Inquiry-based learning process and, therefore, provide additional levels of detail for (re)designing teaching practice.

Finally, the proposed TLA method could be exploited within the research domain of future Cognitive Learning Analytics Systems. In particular, such automated systems are envisaged to be able to learn and interact naturally with people, so as to extend what either man or machine could do on their own; thereby allowing rapid decision-making based on data. In this context, the proposed approach takes a step towards this direction by allowing automated analysis of IED in terms of guidance provided and visualizes the results to teachers to inform their decision making. As future work, new systems should aim to further streamline the processes of designing and evaluating educational designs (across different subject domains and educational levels). For example, such systems could include recommendation capabilities to support teachers during the design and the evaluation (also revision) of their educational designs. In particular, for the design stage, building on the case study presented in this paper, future Cognitive Analytics systems could offer recommendations of guidance tools that could appropriately alleviate shortcomings on the design of IED. Furthermore, regarding the evaluation (and revision) stage, future Cognitive Analytics systems could generate recommendations on guidance tools that could better meet the needs of individual students based on their profiles. This *prescriptive* dimension of TLA methods (i.e., offering recommendations on 'what to do' based on what has already happened) is still very under-researched (Sergis & Sampson, 2017) and additional work needs to be invested in this direction, to effectively support teachers' reflection with usable, actionable insights.

References

Ali, L., Hatala, M., Gašević, D., & Jovanović, J. (2012). A qualitative evaluation of evolution of a learning analytics tool. *Computers & Education, 58*(1), 470-489.

Avramides, K., Hunter, J., Oliver, M., & Luckin, R. (2015). A method for teacher inquiry in cross-curricular projects: Lessons from a case study. *British Journal of Educational Technology, 46*(2), 249-264.

Belland, B. R. (2017). Computer-Based Scaffolding Strategy. In B Belland (Ed), *Instructional Scaffolding in STEM Education* (pp. 107-126). Springer International Publishing.

Belland, B. R., Walker, A. E., Kim, N. J., & Lefler, M. (2017). Synthesizing

results from empirical research on computer-based scaffolding in STEM education: A meta-analysis. *Review of Educational Research, 87*(2), 309-344.

Charleer, S., Klerkx, J., & Duval, E. (2015). Exploring Inquiry-Based Learning Analytics through Interactive Surfaces. In *Proceedings of the 5th International Learning Analytics & Knowledge Conference*.

D'Angelo, C., Rutstein, D., Harris, C., Bernard, R., Borokhovski, E., & Haertel, G. (2014). *Simulations for STEM learning: Systematic review and meta-analysis.* Menlo Park, CA: SRI International.

D'mello, S.K., & Kory, J. (2015). A review and meta-analysis of multimodal affect detection systems. *ACM Computing Surveys, 47*(3), 43.

De Jong, T., Van Joolingen, W. R., Giemza, A., Girault, I., Hoppe, U., Kindermann, J., ... & Weinbrenner, S. (2010). Learning by creating and exchanging objects: The SCY experience. *British Journal of Educational Technology, 41*(6), 909-921.

de Jong, T., & Lazonder, A.W. (2014). The guided discovery principle in multimedia learning. In R.E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 371-390), Cambridge: Cambridge University Press

De Jong, T., Sotiriou, S., & Gillet, D. (2014). Innovations in STEM education: The Go-Lab federation of online labs. *Smart Learning Environments, 1*(1), 1-14.

De Jong, T., Linn, M.C., & Zacharia, Z.C. (2013). Physical and virtual laboratories in science and engineering education. *Science, 340*(6130), 305-308.

Donnelly, D. F., Linn, M.C., & Ludvigsen, S. (2014). Impacts and characteristics of computer-based science inquiry learning environments for precollege students. *Review of Educational Research, 84*(4), 572-608.

Dyckhoff, A. L., Zielke, D., Bültmann, M., Chatti, M. A., & Schroeder, U. (2012). Design and Implementation of a Learning Analytics Toolkit for Teachers. *Educational Technology & Society, 15*(3), 58-76.

Dyckhoff, A.L., Lukarov, V., Muslim, A., Chatti, M.A. & Schroeder, U. (2013). Supporting action research with learning analytics. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 220-229).

Eckhardt, M., Urhahne, D., Conrad, O., & Harms, U. (2013). How effective is instructional support for learning with computer simulations?. *Instructional Science, 41*(1), 105-124.

Freeman, B., Marginson, S., & Tytler, R. (2015). Widening and Deepening the STEM Effect, In B. Freeman, S. Marginson, & R. Tytler (Eds.), *The Age of STEM: Policy and practice in science, technology, engineering and mathematics across the world* (pp. 1 – 21), New York: Routledge.

Gijlers, H., & de Jong, T. (2009). Sharing and confronting propositions in collaborative inquiry learning. *Cognition and Instruction, 27*(3), 239-268.

Gillet, D., De Jong, T., Sotirou, S., & Salzmann, C. (2013). Personalised learning spaces and federated online labs for STEM Education at School. In *Proceeddings. of the 2013 IEEE Global Engineering Education Conference* (pp. 769-773).

Gillet, D., Vozniuk, A., Rodríguez-Triana, M.J., Holzer, A. (In Press). Agile, Versatile, and Comprehensive Social Media Platform for Creating, Sharing, Exploiting, and Archiving Personal Learning Spaces, Artifacts, and Traces. In *Proceedings of The World Engineering Education Forum.*

Gillies, R. M., Nichols, K., Burgh, G., & Haynes, M. (2014). Primary students' scientific reasoning and discourse during cooperative inquiry-based science activities. *International Journal of Educational Research, 63*, 127-140.

Girault, I., & d'Ham, C. (2014). Scaffolding a Complex Task of Experimental Design in Chemistry with a Computer Environment. *Journal of Science Education and Technology, 23*(4), 514-526.

Haya, P. A., Daems, O., Malzahn, N., Castellanos, J., & Hoppe, H. U. (2015). Analysing content and patterns of interaction for improving the learning design of networked learning environments. *British Journal of Educational Technology, 46*(2), 300-316.

Hou, H. T. (2015). Integrating cluster and sequential analysis to explore learners' flow and behavioral patterns in a simulation game with situated-learning context for science courses: a video-based process exploration. *Computers in human behavior, 48*, 424-435.

Hwang, G. J., Wu, P. H., Zhuang, Y. Y., & Huang, Y. M. (2013). Effects of the inquiry-based mobile learning model on the cognitive load and learning achievement of students. *Interactive Learning Environments, 21*(4), 338-354.

Johnson, L., Adams Becker, S., Estrada, V., & Freeman, A. (2015). *NMC Horizon Report: 2015 K-12 Edition.* The New Media Consortium.

Kim, M.C., & Hannafin, M.J. (2011). Scaffolding 6th graders' problem solving in technology-enhanced science classrooms: A qualitative case study. *Instructional Science, 39*(3), 255-282.

Kim, H. J., & Pedersen, S. (2011). Advancing young adolescents' hypothesis-development performance in a computer-supported and problem-based learning environment. *Computers & Education, 57*(2), 1780-1789.

Krüger, A., Merceron, A., & Wolf, B. (2010). A data model to ease analysis and mining of educational data. In *Proceedings of the 3rd International Conference on Educational Data Mining* (pp. 131-140).

Kyza, E. A., Constantinou, C. P., & Spanoudis, G. (2011). Sixth graders' co-construction of explanations of a disturbance in an ecosystem: Exploring relationships between grouping, reflective scaffolding, and evidence-based explanations. *International Journal of Science Education, 33*(18), 2489-2525.

Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning design. *American Behavioral Scientist, 57*(10), 1439-1459.

Lops, P., de Gemmis, M., & Semeraro, G. (2011). Content-Based Recommender Systems: State of The Art and Trends. In F. Ricci, L. Rokach, B. Shapira and P.B Kantor, (Eds.), *Recommender systems handbook* (pp. 73-105), US: Springer.

Manske, S., Chounta, I.A., Rodríguez-Triana, M.J. , Gillet, D., & Hoppe, H.U. (2015). Exploring Deviation in Inquiry Learning: Degrees of Freedom or Source of Problems?. In *Proceedings of the 23rd International Conference on Computers in Education*.

Mikroyannidis, A., Okada, A., Scott, P., Rusman, E., Specht, M., Stefanov, K., ... & Kikis-Papadakis, K. (2013). weSPOT: A Personal and Social Approach to Inquiry-Based Learning. *Journal of Universal Computer Science, 19*(14), 2093-2111.

Monroy, C., Rangel, V. S., & Whitaker, R. (2014). A strategy for incorporating learning analytics into the design and evaluation of a K–12 science curriculum. *Journal of Learning Analytics, 1*(2), 94-125.

Mulholland, P., Anastopoulou, S., Collins, T., Feisst, M., Gaved, M., Kerawalla, L., ... & Wright, M. (2012). nQuire: technological support for personal inquiry learning. *IEEE Transactions on Learning Technologies, 5*(2), 157-169.

Nguyen, Q., Rienties, B., & Toetenel, L. (2017). Unravelling the dynamics of instructional practice: a longitudinal study on learning design and VLE activities. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 168–177).

Pedaste, M., Mäeots, M., Siiman, L. A., De Jong, T., Van Riesen, S. A., Kamp, E. T., ... & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational research review, 14*, 47-61.

Rienties, B., & Toetenel, L. (2016). The impact of learning design on student behaviour, satisfaction and performance: A cross-institutional comparison across 151 modules. *Computers in Human Behavior, 60*, 333-341.

Rodríguez-Triana, M.J., Martínez-Monés, A., Asensio-Pérez, J.I., & Dimitriadis, Y. (2015). Scripting and monitoring meet each other: Aligning learning analytics and learning design to support teachers in orchestrating CSCL situations. *British Journal of Educational Technology, 46*(2), 330-343.

Rodríguez-Triana, M.J., Govaerts, S., Halimi, W., Holzer, A., Salzmann, C., Vozniuk, A., ... & Gillet, D. (2014). Rich open educational resources for personal and inquiry learning: Agile creation, sharing and reuse in educational social media platforms. In *Proceedings of IEEE International Conference on the Web and Open Access to Learning* (pp. 1-6).

Ruiz-Calleja, A., Vega-Gorgojo, G., Asensio-Pérez, J. I., Bote-Lorenzo, M. L., Gómez-Sánchez, E., & Alario-Hoyos, C. (2012). A Linked Data approach for the discovery of educational ICT tools in the Web of Data. *Computers & Education, 59*(3), 952-962.

Sharples, M., Scanlon, E., Ainsworth, S., Anastopoulou, S., Collins, T., Crook, C., ... & O'Malley, C. (2015). Personal inquiry: Orchestrating science investigations within and beyond the classroom. *Journal of the Learning Sciences, 24*(2), 308-341.

Sergis, S. & Sampson, D. (2017). Teaching and Learning Analytics To Support Teacher Inquiry: A Systematic Literature Review. In A Ayala (Ed.), *Learning analytics: Fundaments, applications, and trends: A view of the current state of the art* (pp. 25-63), International Publishing: Springer.

Sergis, S. & Sampson, D.G. (2016a). Towards a Teaching Analytics Tool for supporting reflective educational (re)design in Inquiry-based STEM Education. In *Proceedings. of the 16th IEEE Conference in Advanced Learning Technologies* (pp. 314-318).

Sergis, S., & Sampson, D.G. (2016b). Learning object recommendations for

teachers based on elicited ICT competence profiles. *IEEE Transactions on Learning Technologies, 9*(1), 67-80.

Siemens, G., & Long, P. (2013). Penetrating the Fog: Analytics in Learning and Education. *EDUCAUSE Review, 46*(5).

Slotta, J.D. (2004) The web-based inquiry science environment (WISE): Scaffolding knowledge integration in the science classroom," In M.C. Linn, E.A. Davis and P. Bell (Eds.), *Internet environments for science education* (pp. 203–232), Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Slotta, J.D., Tissenbaum, M., & Lui, M. (2013). Orchestrating of complex inquiry: three roles for learning analytics in a smart classroom infrastructure. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 270-274)

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management, 45*(4), 427-437.

Quintana, C., Reiser, B.J., Davis, E.A., Krajcik, J., Fretz, E., Duncan, R.G., ... & Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *The journal of the learning sciences, 13*(3), 337-386.

van Joolingen, W. R., de Jong, T., Lazonder, A. W., Savelsbergh, E. R., & Manlove, S. (2005). Co-Lab: research and development of an online learning environment for collaborative scientific discovery learning. *Computers in Human Behavior, 21*(4), 671-688.

Van Leeuwen, A., Janssen, J., Erkens, G., & Brekelmans, M. (2014). Supporting teachers in guiding collaborating students: Effects of learning analytics in CSCL. *Computers & Education, 79*, 28-39.

Verbert, K., Govaerts, S., Duval, E., Santos, J. L., Van Assche, F., Parra, G., & Klerkx, J. (2014). Learning dashboards: an overview and future research opportunities. *Personal and Ubiquitous Computing, 18*(6), 1499-1514.

Vozniuk, A., Rodrıguez-Triana, M.J., Holzer, A., Govaerts, S., Sandoz, D., & Gillet, D. (2015). Contextual Learning Analytics Apps to Create Awareness in Blended Inquiry Learning. In *Proceedings of the 14th IEEE International Conference on Information Technology Based Higher Education and Training* (pp. 1-5).

Vozniuk, A., Rodríguez-Triana, M., Holzer, A., Gillet. D. (2016) Combining Content Analytics and Activity Tracking to Identify User Interests and Enable Knowledge Discovery. In *Proceedings of the 6ᵗʰ International Workshop on Personalization Approaches in Learning Environments* (pp. 1-7)

Wang, H. Y., Duh, H. B. L., Li, N., Lin, T. J., & Tsai, C. C. (2014). An investigation of university students' collaborative inquiry learning behaviors in an augmented reality simulation and a traditional simulation. *Journal of Science Education and Technology, 23*(5), 682-691.

Wasson, B., Hanson, C., & Mor, Y. (2016). Grand Challenge Problem 11: Empowering Teachers with Student Data. In J. Eberle, K. Lund, P. Tchounikine and F. Fischer (Eds.), *Grand Challenge Problems in Technology-Enhanced Learning II: MOOCs and Beyond* (pp. 55-58) International Publishing: Springer.

White, B., Frederiksen, J., Frederiksen, T., Eslinger, E., Loper, S., & Collins, A. (2002). Inquiry Island: Affordances of a multi-agent environment for scientific inquiry and reflective learning. In *Proceedings of the Fifth International Conference of the Learning Sciences* (pp. 1-12).

Wu, H. L., & Pedersen, S. (2011). Integrating computer-and teacher-based scaffolds in science inquiry. *Computers & Education, 57*(4), 2352-2363.

Yang, T. C., Chen, S. Y., & Hwang, G. J. (2015). The influences of a two-tier test strategy on student learning: A lag sequential analysis approach. *Computers & Education, 82*, 366-377.

Zacharia, Z. C., Manoli, C., Xenofontos, N., de Jong, T., Pedaste, M., van Riesen, S. A., ... & Tsourlidaki, E. (2015). Identifying potential types of guidance for supporting student inquiry when using virtual and remote labs in science: a literature review. *Educational technology research and development, 63*(2), 257-302.

Zervas, P., & Sampson, D. (In Press). Supporting Reflective Lesson Planning based on Inquiry Learning Analytics for Facilitating Students' Problem Solving Competence Development: The Inspiring Science Education Tools. In R. Huang, N.–S. Chen and Kinshuk (Eds.), *Authentic Learning through Advances in Technologies*, Springer.

Zervas, P., Sergis, S., Sampson, D. G., & Fyskilis, S. (2015). Towards competence-based learning design driven remote and virtual labs recommendations for science teachers. *Technology, Knowledge and Learning, 20*(2), 185-199.

Acknowledgments