Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning

Katerina Mangaroska, Michail Giannakos, *Senior Member, IEEE*

Abstract—As the fields of learning analytics and learning design mature, the convergence and synergies between these two fields became an important area for research. This paper intends to summarize the main outcomes of a systematic literature review of empirical evidence on learning analytics for learning design. Moreover, this paper presents an overview of what and how learning analytics have been used to inform learning design decisions and in what contexts. The search was performed in seven academic databases, resulting in 43 papers included in the main analysis. The results from the review depict the ongoing design patterns and learning phenomena that emerged from the synergy that learning analytics and learning design impose on the current status of learning technologies. Finally, this review stresses that future research should consider developing a framework on how to capture and systematize learning design data grounded in learning analytics and learning theory, and document what learning design choices made by educators influence subsequent learning activities and performances over time.

Keywords—Learning analytics, Learning design, Empirical studies, Systematic literature review.

I. INTRODUCTION

The use of analytics to discover important learning phenomena (e.g. moment of learning or misconception) and portray learners' experiences and behaviors, is evident and commonly accepted due to the pervasiveness of learning technologies. Learning analytics holds a critical role in understanding human learning, teaching, and education, by identifying and validating relevant measures of processes, outcomes, and activities. In addition, learning analytics supports and promotes evidence-based practices derived from evaluation and assessment of learners' progress, motivation, attitudes, and satisfaction. However, learning analytics lacks theoretical orientation that can assist researchers to explain inconsistencies, avoid misinterpretations, and consider and clarify any contextual conditions (e.g. instructional, sociological, psychological, etc.) that affect learning [1], [2]. Moreover, Reimann highlights that "atheoretical approaches to learning analytics might produce misconceptions because it is the logical (and ethical) error of using descriptions of the past as prescriptions for the future" [2, p.136]. Consequently, without theoretical grounding of learning analytics, and contextual interpretation of the collected data, learning analytics design capabilities are limited. From this perspective, learning design is utterly important as it provides the framework for analyzing and interpreting data, learner's behavior, and successful or inefficient learning patterns.

Learning design defines the educational objectives and the pedagogical approaches that educators can reflect upon, take decisions, and make improvements. In other words, learning design is the "application of methods, resources and theoretical frameworks to achieve a particular pedagogical goal in a given context" [3, p.88]. Moreover, learning design "documents the sequence of learning tasks and teaching methods" as main premises for re-usability and transferability of good practices across educational contexts [4, p.3]. Yet, past research was focused on "conceptualizing learning design principles, without evaluating what happens after the design process" [5, p.333]. In addition, several studies have tried to understand and improve the learning design experiences by utilizing learning analytics, but only few of them establish the usage of learning analytics on existing principles and theories in learning sciences, educational research, technology acceptance, and human-computer interaction [6], [7].

As it can be observed from the literature, learning analytics is an interdisciplinary field embracing methods and approaches from various disciplines, and as such, lacks a consolidated model to systematize how those disciplines are merged together [1]. Moreover, research is missing to measure what learning design decisions affect learning behavior and stimulate productive learning environment, as well as, what learning analytics generate actionable design insights for various groups of stakeholders [1]. To bridge the gap, this paper centers in a systematic literature review with an aim to examine the intersection between learning analytics and learning design, and provide important insights beyond the specific research findings within the individual disciplines. Although the field of learning analytics is still relatively young [8], as indicated by Google Trends too (see Figure 1), enough work has already been done to conduct a review [9], [4]. Thus, the study addresses the following research questions:

RQ1: What is the current status of learning analytics for learning design research, seen through the lens of educational contexts (i.e. users and rational for use), distribution of pedagogical practices, and methodologies (i.e. types of data and data analysis techniques employed)?

RQ2: What learning analytics have been used to inform learning design decisions, and to what extent learning analytics have the capacity to support dynamic and data-driven learning design decisions?
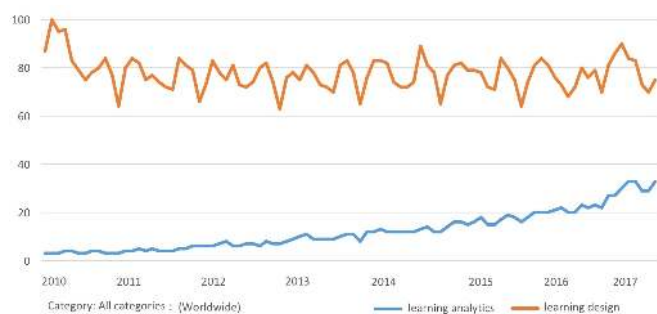
Fig. 1. Search interest in Learning Analytics (blue line) and Learning Design (orange line) according to Google Trends.

The rest of the paper is organized as follows. In the next section the authors present the related work; the third section describes the methodology used for the literature review describing how the studies were selected and analyzed. The fourth section presents the research findings derived from the data analysis based on the specific areas of focus. Finally, in the last section, the authors discuss the results and identify gaps, while making suggestions for future considerations.

II. RELATED WORK

A. Learning analytics

In the last ten years, learning analytics highlighted the gradual shift from technological towards educational perspective, despite its roots in business intelligence, recommender systems, and educational data mining [10]. Its emergence as a separate field is due to the increasing trend of digitization in the field of education, the appearance of distributed learning environments, and the increased engagement in online learning experiences [11]. The practice of learning analytics evolved around the idea of harnessing the power of digital technologies to collect traces that learners leave behind, in order to understand activities and behaviors associated with their learning [8]. As a result, learning analytics holds the potential to: 1) explain unexpected learning behaviors, 2) identify successful learning patterns, 3) detect misconceptions and misplaced effort, 4) introduce appropriate interventions, and 5) increase users' awareness of their own actions and progress [11]. Undoubtedly, learning analytics is an interdisciplinary field that embraces a holistic approach to study learning contexts and environments, and to address questions in educational research [12]. As a term it has a generally accepted definition, adopted by the Society for Learning Analytics Research (SoLAR): "Learning analytics is the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs" [13].

Previous research [14], supports and promotes evidence-based practices of learning analytics potential in understanding and optimizing the complexities of the learning process. However, learning analytics should not promote "one-size-fits-all" research due to the fact that students' individual differences have strong implications on the learning behavior, motivation,

and engagement [15], [16]. Consequently, designing and applying personalized learning activities could improve student's performance and learning progress [17]. Along the same lines, Papamitsiou and Economides [9], [18] systematized the significant contribution of learning analytics empirical research and identified some early indications of how learning analytics and data mining might support personalized and adaptive learning experiences utilizing rich data. Consequently, developing and employing personalized learning and feedback mechanisms to support learners to follow and regulate their progress, involves more than just data easily collected. It actually tackles the learning design activities grounded in theory and data, where educators decide how to introduce the analytics and how to frame aspects of their interpretation within a socio-technical system. This is due to the fact that "human decision-making and consequent actions are as much a part of any successful analytics solution as its technical components" [19, p.4].

B. Learning design

Learning design is another field associated with online and technology enhanced learning (TEL) research in the late 1990s and 2000s, that holds different theoretical background than the domain of instructional design [3]. The term was coined to replace the already established term instructional design based on behaviourism and cognitivism [20], [21] and to include educational interventions based on socio-constructivist approaches mediated by technology [3], [22]. However, the field of learning designed emerged from the perception of educator's role in education, as Laurillard phrased it: "not to transmit knowledge to a passive recipient, but to structure the learner's engagement with the knowledge, practising the high-level cognitive skills that enable them to make that knowledge their own" [23, p.527].

At present, learning design is very diverse, because the way the term is conceptualized is contingent on observer's choice of perspective [24], [3]. This is representative of the emergent discourse among the researchers and practitioners that shape the field of learning design [25]. Nonetheless, learning design must be conceptualized before it can be utilized as a process that leads to explicit and sharable design outputs for learning [26]. Thus, some researchers see learning design as "a form of documentation of pedagogical intent that provides the context to interpret the analytics from the diverse data sets" [4, p.1]. For others, learning design is "a methodology that educators use and communicate with each other to make informed decisions in designing learning activities and interventions with effective use of resources and technologies" [25, p.121]. However, in a more general sense, learning design can be defined as "the description of the teaching-learning process that takes place in a unit of learning (e.g., a course, a lesson or any other designed learning event)." [27, p.14].

Although learning design and instructional design perspectives have a substantial overlap in the literature, learning design emphasizes more the learner's context and the constructivist approach in the learning activities [3]. Thus, learning design can be seen "as a creative process of perpetual educational innovation grounded in well-defined context of practice and

pedagogical theory for generating new solutions to various educational challenges." [3, p.93]. On one hand, its aim as a field is to explore issues in education and support educators to make pedagogically grounded decisions in their teaching practices [25]. However, a common language for learning design is needed, in order to streamline the process of constructing, validating, and disseminating design knowledge among the research community. This way learning design could have a positive impact on sharing, discussing, and improving the educational processes [22]. This is also the idea behind Perisco and Pozzi call for multi-dimensional framework drawing together a number of approaches and tools for design of learning, rather than just schemes and graphical representations [28].

On the other hand, educators need to orchestrate all of the activities that learning design encompasses, including the constraints and challenges (e.g. time, attention, cognitive overload) they face in their everyday practice [29], [30]. Orchestration is a research phenomenon that deals with the complexity of learning design activities and application of technological innovations in education [31]. As a definition, orchestration covers "the process by which teachers and other actors design, manage, adapt and assess learning activities, aligning the resources at their disposal to achieve the maximum learning effect, informed by theory while complying pragmatically with the contextual constraints of the setting" [32]. Moreover, the field of orchestration research proposes tools and frameworks to conceptualize and adapt the available pedagogical and technological innovations, as a way to achieve improvement in teaching and learning. As such, it cannot be overlooked in the discourse apropos of learning design and learning analytics [33].

C. Learning design taxonomy

The term learning design in this paper refers to the process of designing effective learning experiences with use of technological innovations and resources. If effective, this process could be shared between educators and reused or adapted. Thus, there are several initiatives to create descriptive framework of instructional practices so that teaching approaches are shared among educators [34]. One of those initiatives is the Open University Learning Design Initiative, that categorize learning design in seven broad learning design activities [34]. *Assimilative* are learning activities in which students attend to information as required by their instructors; *Finding and handling information* includes learning activities which focus on skills development; *Communicative* activities encompass all communication between students, or students and instructors; *Productive* activities focus on active learning where students build artifacts; *Experiential* activities support students to apply their knowledge in real-world settings; *Interactive/adaptive* activities include role-play, problem-based scenarios in simulated experiments; and finally *assessment* activities that include all forms of assessment.

Having understanding of specific issues and phenomena during the implementation of learning design activities in technology-rich educational settings is of utmost importance [31]; and in the same time very challenging to be addressed without the support of learning analytics. Moreover, there is no other classification of learning design concepts, that has been empirically used to compare, on a large scale, module designs across disciplines in university institutions [5]. Consequently, the authors want to explore what type of learning design activities have been used in the selected studies, and what learning analytics the researchers have applied to see how students' behavior relates to a specific learning activity. In particular, the focus is towards well-designed learning design activities that provide foundation for effective scaffolding of student's learning behavior.

D. Theoretical alignment

Adoption of data-driven approaches in learning analytics emphasize the power of data science methods with unprecedented amounts of data collected from students and teachers in distributed learning environments [35]. However, data-driven approaches were later recognized as not sufficiently informative [36]. Furthermore, as highlighted by Gašević et. al, [36] one of the most important tasks of learning analytics is the "development of measures that can increase the understanding into the learning processes and interpret those measures to inform existing theories for the purpose of developing actionable initiatives in teaching practices and design of learning environments". Consequently, theory orientation in learning analytics is essential. It helps to identify meaningful patterns and associations between digital traces and learning outcomes [37], [1]; to decide what questions to research to improve TEL [38]; what methods and analysis to select [39]; and how to interpret the outcomes to produce actionable insights for various stakeholders [36]. In the existing literature, there is a reference model that identifies four critical learning analytics dimensions: what (data is gathered, managed and analyzed), who (is the target audience), why (data is gathered and analyzed, and how (data will be analyzed) that need to be considered when designing learning activities [40]. Similarly to Chatti et al. [40] Greller and Drachsler [41] identified six critical dimensions of leaning analytics that need to be covered by the design to ensure use of learning analytics in an "educationally beneficial way". Another conceptual four dimensional framework proposed by Martinez et al. [42] provides guidelines how to design learning analytics technologies that will address orchestration challenges utilizing data from interactive surfaces. Finally, a new conceptual framework (Orchestrating Learning Analytics - OrLA) is proposed to overcome the gaps in the adoption of learning analytics innovations by supporting inter-stakeholder dialogue at the practitioner level [33].

Despite the widely accepted and used term "learning analytics", the reference to "learning" is still young as learning analytics only recently began to make connection with learning theories [10], [15]. This deficiency leads to another current issue of misalignment between the information generated by learning analytics, and the needs, problems, and concerns teachers have with the learning design activities. The reason for this misalignment can also be found in the gap between data easily captured from system logs and data that is pedagogically valuable. One possible solution to overcome the disconnec-

tion between research and everyday pedagogical practice, is development of a common framework that will systematize the process of establishing effective solutions using learning analytics grounded in theory to open educational issues. The advantage of developing a common framework should be seen in establishing understanding, validity, reliability, and direct support by clear guidance of the types of analytics and tools essential for particular learning contexts. For example, we have Perisco and Pozzi [28] framework of representations, approaches, and tools from which teacher's training in learning design can draw upon, combined with hands on experience. Next is the proposed conceptual framework that links learning analytics to learning design with the aim to support enquiry-based evaluation and scaffolding of learning designs [43]. Furthermore, the concept of orchestrating learning analytics (OrLA) aims to provide conceptual framework and guidelines that support teacher's complex activities utilizing learning analytics in authentic educational practices [33]. Another possible solution to overcome the disconnection between research and everyday pedagogical practice is utilizing effective learning techniques grounded in theory to help students achieve their learning goals that can later be empirically validated and modeled to more directly guide behavior [44].

Although learning analytics is receiving close attention in the TEL community, there are issues that the field is struggling to answer. This is due to the lack of theoretical grounding in interdisciplinary approaches, such as educational research, learning sciences, psychology, human-computer interaction, data mining, and research methods [1]. Therefore, Gašević et al. [1] proposed a consolidated model how theory, design, and data mutually interact and inform decisions related to practice, privacy, ethics, policy, and standards. This model incorporates *theory* to identify which associations between learning analytics and learning outcomes are meaningful, and as such, insert them into analytical models; *design* grounded in learning theory and tailored to activate particular learning mechanisms; and *data* to identify indicators and measures in learning analytics far from just using counts of click-streams data.

E. Learning analytics for learning design

As the field of learning analytics matures, its convergence and synergy with the field of learning design becomes an important area for research. The alignment between learning analytics and learning design derives from the possibility to:

1) utilize learning analytics to "facilitate the drive from tacit educational practice to explicit" [34],
2) utilize learning design in pedagogical context to translate the learning analytics findings into meaningful information [4], [28], [45].

Rienties et al. [34] presented a review study of ten years research at Open University UK in aligning learning design with learning analytics, underlining the importance of learning design in learning experiences and teaching practices. Moreover, they also emphasized that "learning design focuses on what students do as part of their learning, rather than on the content that is delivered by the teacher" [46]. However, there is a paucity of evidence for how learners respond to different learning designs, that hinders researchers to explore which pedagogies and conceptualizations work best [34]. Furthermore, several studies highlighted and acknowledged the need to align both approaches with a conceptual framework that will facilitate further maturation of the fields [36], [6], [47], [48], [49]. Having a generally accepted framework might help researchers to understand how specific design elements (i.e. design of learning tasks) influence students' behaviors, engagement, and learning; while at the same time discover how students engage and learn within an authentic pedagogical and technological context. In addition, Lockyer and Dawson [50] work complements these studies by demonstrating the evaluative potential of learning analytics to inform pedagogical action, and accordingly, improvements in the learning design. In their later research, they highlighted the importance of interpreting learners' behaviors and reactions, as well as developing a conceptual model for educators' use of learning analytics when developing learning design strategies [4]. Additionally, Persico and Pozzi [28] collocated much of the work done in the fields of learning analytics and learning design, and highlighted the issues using a multitude of different approaches, tools, and representations that are not interoperable, and as such present various epistemological issues. Finally, aligning the both approaches with a conceptual framework could also increase the communication among the various stakeholders about the adoption of learning analytics at practitioner level [33].

Although learning analytics and learning design share common goals, their alignment and convergence is still limited. To address this issue, the research community needs to reach out in both directions; learning analytics needs to consider educational research and theory in the design of analytics; while learning design needs to utilize data mining and information contextualization before designing for analytics use [1].

Consequently, for the purpose of this review study and interpretation of the results based on the findings from the wide literature, the authors outline a definition of learning analytics for learning design in the context of the proposed research questions, as: *"usage of learners and educators-produced data to discover behavior patterns that are of core interest to both groups, for the purpose of devising explicit, sharable, and reusable learning designs, practices, resources, and tools, aimed at achieving educational goals in a given learning context"*. Considering this definition, the current review study investigates what empirically-based learning analytics are employed to inform actionable design decisions using what types of data and analysis methods; as well as to report the influence of those analytics-driven design-decisions in learning and teaching.

III. METHODOLOGY

To answer the research questions, the authors decided to conduct a systematic review of the literature by following transparent procedure adopted in the field of computer science in order to minimize potential researcher biases and support reproducibility [51].

TABLE I.   INCLUSION/EXCLUSION (I/E) CRITERIA

Inclusion criteria	Exclusion criteria
The research addresses educational practices.	No abstract.
The research is an empirical study.	The paper is written before 2010.
The research explores data analytics role in supporting learning design activities.	The paper is not written in English.
Research target audience are students and/or teachers.	Not a research/peer-reviewed paper (e.g. editorial, workshop, expert opinion, work-in-progress).

A. Systematic review planning

To the authors knowledge, no previous work aimed at producing a systematic and comprehensive overview of the existing empirical work on the convergence and synergy between learning analytics and learning design. Thus, the aim of this paper is to systematize and summarize the empirical work in the field over time, and aggregate the insights from the review. The comprehensive review provided in this paper could help different stakeholders (especially instructional designers and TEL researchers) to understand what has already been explored, implemented, and validated at the intersection of learning analytics and learning design. In particular, the authors aim to investigate the current status of learning analytics for learning design; classify what learning analytics indicators have been used to inform learning design decisions; and offer a synthesis of the existing approaches towards the alignment of learning design and learning analytics.

Search strategies. To find primary studies relevant for this review study, the authors decided to include only empirical peer-reviewed work as a standard for the quality of the selected studies. The peer-reviewed papers need to be published in one of the main five academic electronic databases in Technology Enhanced Learning (TEL): ACM DL, IEEE Explore, Springer-Link, Science Direct, and Wiley, and two additional databases, SAGE and ERIC. The second cycle included an independent search in key educational technology journals listed in the Google metrics sub-category: Educational Technology; i.e. Computers & Education, British Journal of Educational Technology (BJET), The Internet and Higher Education, Journal of Educational Technology & Society, Journal of Computer Assisted Learning, Educational Technology Research and Development, International Journal of Computer-Supported Collaborative Learning, IEEE Transactions on Learning Technologies, and the International Conference of Learning Analytics and Knowledge (LAK). Moreover, a search in Google Scholar for potentially relevant literature that is not normally indexed in the most common academic databases (e.g., ICLS, EDM, CSCL etc.) was also performed. The third and final cycle included a search in the reference section for each selected paper in order to find additional relevant papers (i.e. the snowball technique).

Selection criteria. The primary studies retrieved from the databases or the educational and technology journals, need to be filtered using different sets of criteria. Initially, the authors will consider four inclusion and four exclusion criteria to select papers to be further analyzed in the review study, as shown in Table I. Next, from the initial selection of papers, the authors will continue the selection process according to another set

of eight quality criteria shown in Table II. These criteria were informed by the proposed Critical Appraisal Skills Programme (CASP) [52], [53] and by principles of good practice for conducting empirical research in software engineering [54]. As Dybå and Dingsøyr [52] specified, the quality criteria needs to cover three main issues (i.e. rigour, credibility, and relevance) that needs to be considered when evaluating the quality of the selected studies. Finally, the retrieved papers can be duplicates, overlapping or extended versions from the same authors. In such cases, the duplicate papers will be immediately discarded, the overlapping papers will be integrated and treated as one paper, while for the extended papers a selection will always be made for the extended publication due to the details the extended version provides.

B. Search string construction

The search string used during the search covers three main terms (analytics, design, and learning) which have to appear in the potentially relevant primary studies. The combination of the three main terms should capture a large scale of potential research at the intersection between learning analytics and learning design. The terms, *analytics* and *design*, are the main topics of the study. However, the authors are only interested how these two terms are used in the field of education; thus adding the third term *learning*. The search string used is: *"analytics"* AND *"design"* AND *"learning"*. Due to the high number of irrelevant papers (i.e., false positives) returned back using the search string *"analytics"* AND *"design"* AND *"learning"*, the authors decided to narrow the search by combining the three words into (*"learning analytics"* AND *"design"*) and (*"analytics"* AND *"learning design"*). In addition to the three main terms, the authors decided to add one more term, *orchestration*, as shown in Figure 2, that is already embraced by the educational researchers, to explain the practical issues and task that are "not directly linked with learning but can shape learning", making it relevant for utilizing learning analytics in learning settings [55]. The authors add this fourth term to capture potential literature that uses the expression orchestration to refer to the complexity of learning design activities not only in the classroom, but also in online or blended learning scenarios, that otherwise might have been omitted. The additional search string used is: *"analytics"* AND *"design"* AND *"learning"* AND *"orchestration"*. Consequently, the authors decided to use the following search strings:

1) *"learning analytics"* AND *"design"*
2) *"analytics"* AND *"learning design"*
3) *"analytics"* AND *"design"* AND *"learning"* AND *"orchestration"*

customized to the specific syntax for each database.

C. Systematic review execution

The first step after constructing the search string includes execution of the search queries in the selected databases and journals, from mid-October to mid-December 2016. One of the researchers searched the titles, abstracts, and keywords of the articles in the included electronic databases, and the
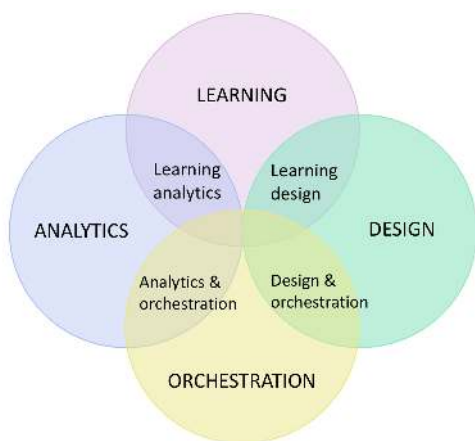
Fig. 2. Search query based on combination of four words

TABLE II.   QUALITY CRITERIA

1. Does the study clearly address the research problem?
2. Is there a clear statement of the aims of the research?
3. Is there an adequate description of the context in which the research was carried out?
4. Was the research design appropriate to address the aims of the research?
5. Does the study clearly determine the research methods (subjects, instruments, data collection, data analysis)?
6. Was the data analysis sufficiently rigorous?
7. Is there a clear statement of findings?
8. Is the study of value for research or practice?

educational and technology journals. A temporal filter was applied since learning analytics is a relatively new field that emerged back in 2010. This search strategy resulted in a total of 3251 "hits" that included 2347 distinct papers, as shown in Table III. In the second step, both researchers (i.e. the authors of this paper) went through the titles, abstracts, metadata, and keywords, of all studies that resulted from step one, to determine their relevance for the systematic review. At this stage, the researchers excluded studies that were not about educational practices or have nothing to do with learning and teaching. For example, the search returned papers about music due to the inclusion of the term *orchestration*. In this step, the researchers followed the four inclusion and four exclusion criteria mentioned in Table I. Moreover, in this stage, the researchers faced two issues. One, there were cases in which some authors used witty titles that could mislead the actual content of the paper. Second, some abstracts were missing, poor, or misleading. Therefore, at this stage, the researchers scanned the full text of those studies, looking at the methodology section and the reported findings. This step returned 288 papers, as shown in Table III. In the third step, each of the 288 studies was assessed independently by both authors, and critically appraised according to the eight criteria shown in Table II. These criteria were informed by CASP and adapted for the purpose of this study following the Quality Assessment form used in a systematic review study about empirical studies of agile software development [52]. Each of the eight criteria was graded on a "yes" or "no" scale. Thus, this step returned 38 papers for which we could say with a

TABLE III.   SEARCH RESULTS BY SOURCE

Source	Raw results	I/E criteria
SpringerLink	473	39
Wiley	258	35
ACM Digital Library	470	40
IEEE Xplore	452	65
Science Direct	306	68
SAGE	108	8
ERIC	280	33
Total	**2347**	**288**

confidence, that the selected studies could make a valuable contribution to this review.

Since almost all of the key educational technology journals mentioned in the systematic review planning section are included in the selected databases, the second search cycle (i.e. independent search in key educational and technology journals), returned no additional papers that needed to be included. The third and final cycle, a search in the reference section of each of the selected 38 papers, also returned no new additional papers to be included in the systematic review analysis.

As of June 2018, the authors performed an additional search (i.e., for 2017) following the same steps for papers published after the initial search period (i.e. 2010-2016). The additional search returned 5 papers.

In conclusion, the search process uncovered a total of 43 papers that were read it entirely, coded, and critically assessed to the review context of this systematic study. The joint probability of agreement measure was 80%, meaning that 80% of the time the two researchers had an overall agreement rate during the selection stages. The two researchers resolved any disagreement in consensus meetings. A summary for the systematic review execution process is shown in Figure 3.
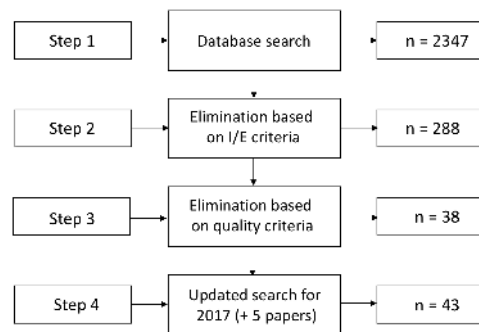

Fig. 3. Summary of the systematic review execution process

D. Data coding

During the coding process, the authors extracted data for more than 15 variables. However, a consensus was reached based on the most important variables that could direct unbiased and ethical analysis of the selected papers, with the final aim to answer the research questions. Thus, the focus areas used for the analysis are: 1) the research design employing the classification presented in [56] and the topic of the study; 2) the

educational context in which the study took place [1] proposed by [9] and the learning scenario; 3) the sample size and the unit of analysis in the study; 4) the pedagogical practices and goals that have been adopted and used (e.g. collaborative learning, self-regulation); 5) the type of the learning platform; 6) the technology and tools used by the subjects during the study (e.g. digital: social media tools, or traditional: mouse, keyboards, pen and paper); 7) the data sources and data collection instruments (e.g. interviews, surveys); 8) the methodology and data analysis techniques; 9) the research objectives reported in the study (e.g. behavior modelling, student assessment), and 10) the impact of learning analytics on subject's behavior and learning performance.

Finally, the authors also strove to understand if the studies integrated Campbell and Oblinger's five step model of learning analytics [57], or in other words, if the studies managed to close the learning analytics loop effectively [58]. Based on the categories and subcategories defined (see Table IV), the two researchers coded all the papers and solved any potential differences. After the coding of the papers, a descriptive analysis of the results was performed to explore the current status of learning analytics for learning design, and classify what learning analytics indicators have been used to inform learning design decisions. The selected variables with a description and a scoring criteria are reported in Table IV. The results from the coding process are reported in Appendix A.

E. Categorization scheme

In order to provide a more holistic view of the current status of learning analytics for learning design, the authors decided to classify what learning analytics metrics have been used, referring to and applying a categorization scheme proposed by [60]. This way, the descriptive analysis will be complemented with categorization of learning analytics according to five perspectives and six data sources. The perspective category includes: 1) individual student (i.e., indicators dedicated to individual student activities; e.g. receptive activities vs active participation indicators), 2) group (i.e., indicators related to a group of students), 3) course (i.e., indicators for monitoring and analyzing the overall course data), 4) content (i.e., indicators that present students' interactions with the learning content), and 5) teacher (i.e., indicators about teacher's actions and activities). The data source category includes: 1) student-generated data, 2) context/local data (i.e., data that surround the student, such as local or mobile data), 3) academic profile (e.g. demographic data, data about past performances), 4) evaluation data, 5) course-related performance, and 6) course meta-data (i.e., data regarding the course structure, goals, resources).

---

[1]VLEs/LMSs: controlled environment, used for gathering learner and activity data, MOOC/social learning: informal, social learning setting, Web-based education: web-based e-learning environments except from VLEs, LMSs and MOOCs, Cognitive tutors: special software, utilized for the needs of the study, Computer-based education: other environments that include some type of computer technology (e.g. desktop applications, etc.) except from those belonging to one of the other categories, Multimodality: learner data in different modalities, Mobility: mobile devices used as the primary learning mediator.

F. Towards learning analytics for learning design taxonomy

In order to comprehensively study learning analytics metrics and learning design decisions, researchers need to find some systematic way to organize, describe, and communicate the research findings using the ontology of the domain (i.e. specifications of conceptualizations) [61]. Thus, selecting and labeling instances under study, and classifying those instances in terms of similarities and differences, leads towards hierarchical classifications of entities within a specific domain [62]. A good taxonomy should separate the entities "into mutually exclusive, unambiguous groups and subgroups that, taken together, include all possibilities" [62, p.52].

Consequently, the authors want to propose a conceptual model towards a learning analytics for learning design taxonomy, deriving classification from existing research and from the review study. The authors will use the already established learning design taxonomy proposed by Rienties et al. [34] without any alterations, and build upon Campbell and Oblinger's [57] five-step model of learning analytics: capture, report, predict, act, and refine. Thus, the mapping between Campbell and Oblinger's five-step learning analytics model and the results from the review study are the following:

1) *capture* will incorporate the different data collection methods commonly used in the selected studies to gather user data;
2) *report* will refer to the techniques researchers used to report the analytics back to the users;
3) *predict* will include the purpose for usage of predictive modelling;
4) *act* will include the actions researches applied;
5) *refine* will refer to the interventions and redesign of learning activities reported in the selected studies.

The conceptual model of learning analytics for learning design means that a taxonomy of learning analytics metrics must be related to the taxonomy of learning design activities for classifying what type of metrics were used for what learning design activities, and what was the outcome in the defined context. This could add more precision and common understanding at the intersection of learning analytics and learning design research. However, establishing selection criteria for extracting pedagogically valuable learning analytics for the development of a taxonomy for research purposes is a challenging task. Therefore, the authors will only propose a conceptual model towards a learning analytics for learning design taxonomy, that will serve as a springboard for further research.

IV. FINDINGS

Analysis of the studies was performed using non-statistical methods considering the variables reported in Table IV. Before continuing with the reporting of the findings, it should be noted that most of the studies had more than one sample population, used more than one data analysis technique, or reported more than one research objective, especially studies that outline two or more case studies. Thus these are aggregated numbers of studies that reported such data. The following findings give an answer to the first research question.

TABLE IV.   Coding schema for the selected research papers

Variable	Description	Scoring criteria
Category [56]	What is the design of the study?	Exp - Experiment, CaseSt - Case study, SDA - Secondary Data Analysis, Etno - Ethnography
Research topic	What is the domain subject?	STEM, CS - Computer Science, SS - Social Sciences, AH - Arts & Humanities
Learning Environment [9]	What is the setting of the learning environment?	VLEs / LMSs - Virtual Learning Environments / Learning Management Systems, MOOC / social learning WBE - Web-based education, CT - Cognitive Tutors, CBE - Computer-based education, MM - Multimodality, Mob - Mobility
Learning scenario	Is it intentional and structured learning provided by an educational institution?	F - Formal, NF - Non-formal, IF - Informal
Population	Sample population	MS - Middle School students, HS - High School students, UG - Undergraduate students, G - Graduate students, E - Educators (e.g. teachers, instructors), R - Researchers
Sample size	Size of sample population	Report actual number of subjects or leave it blank
Unit of analysis	What is the entity that is analyzed in the study?	I - Individual, T - Team (or group), C - Course
Pedagogical approach [59]	What pedagogical approach is adopted?	PBL - Problem-based learning, SRL - Self-regulated learning, IBL - Inquiry-based learning, GBL - Game-based learning, CSCL - Computer Supported Collaborative Learning Constructivism
Learning platform	What type of learning management system is used?	Moodle, Blackboard, Other - write down if reported
Technology and tools	What type of tools and technology are being used by the subjects in the study?	Write the reported technologies and tools or leave it blank. See Appendix A.
Data collection [56]	Type of data source/collection methods	Write the reported data collection methods or leave it blank. See Appendix A.
Methodology	Type of methodology used	Qual - Qualitative, Quant - Quantitative, MMs - Mixed methods
Data analysis [56]	What type of data analysis methods have been used?	DS - Descriptive statistics IS - Inferential statistics: P - Parametric NP - Non-parametric
Research objective	What has been examined in the study?	Write down if authors reported research objectives or leave it blank. See Appendix A.
Behavior	What was the impact of learning analytics on subject's behavior?	Write down if authors reported or leave it blank. See Appendix A.
Performance	What was the impact of learning analytics on learning performance?	Write down if authors reported or leave it blank. See Appendix A.

Publication and study design. In regard to the journal/conference of publication, most of the studies are published in one acknowledged peer-reviewed journal - BJET, and one acknowledged peer-reviewed conference - LAK, as shown in Table V. The high number of published papers in BJET was due to a special issue named "Teacher-led Inquiry and Learning Design" in 2015. The published work shows an expanding interest of exploring the intersection between learning analytics and learning design in the last two years. Although it is still in its infancy, one can see that it receives recognition in the research community. Moreover, the increasing trend shown by years in Fig. 4 also indicates the expanding interest of exploring the domain of learning analytics for learning design.

When it comes to distribution of the selected studies according to the adopted research strategy [56], majority of the papers were case studies (n = 36 studies), following by experiments (n = 5 studies), ethnography (n = 1 studies), and secondary data analysis (n = 1 studies). Regarding the research topic (as reported in the selected studies), the dominant subjects come from computer science (n = 12 studies), of which 6 studies were in programming courses; following STEM (n = 8 studies) of which 3 studies were concentrating on mathematics and statistics; social sciences (n = 8 studies), and arts and humanities (n = 3 studies).

Sample population and unit of analysis. The predominant sample population in the selected papers consisted of under-
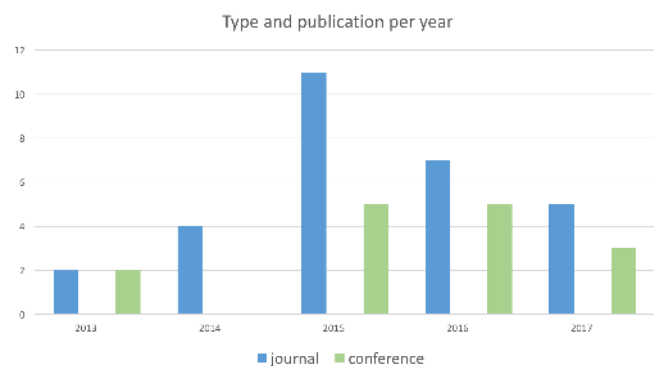
Fig. 4.   Distribution of published work by type and number of publications per year

TABLE V.   DISTRIBUTION OF PUBLISHED WORK BY
JOURNAL/CONFERENCE

Journal/Conference	Num. of studies
British Journal of Educational Technology	6
Computers and Education	1
Computers in Human Behavior	4
The Internet and Higher Education	2
Entertainment Computing	1
American Behavioral Scientist	1
LAK	8
IEEE Transactions on Learning Technologies	3
International Journal of CSCL	1
Journal of Learning Analytics	1
Journal of Computer Assisted Learning	2
Other	13

graduate students (n = 19 studies) and educators (n = 19 studies) as shown in Table VI. Some studies reported only students as a sample population (n = 13 studies) without referring to a specific category, and only two studies included PhD students as a sample population.

Almost all of the studies reported the sample size, that ranged from 4 to 111.256 learners. Substantial sample size was reported only from studies that examined large-scaled MOOCs. Therefore, the researchers decided to calculate the median and the mode of the samples (not considering the sample size reported from MOOCs). As a result, for learners (including middle school, high school, undergraduate and graduate students) the median is 40, and the mode is 15. For educators, the median is 7, and the mode is 12.

On the other hand, the unit of analysis is a very important indicator which defines what type of data should be collected and from whom [56]. Thus, the largest number of papers (n = 12 studies) reported group as their unit of analysis, individuals (n = 9 studies), and pairs (n = 1 study). Only one study reported the course as a unit of analysis to examine the student learning behavior without ignoring the instructor's role or the interaction between the students and the instructor [63].

Setting of the learning environment. Based on the learning settings, most studies were conducted within VLEs/LMSs (n = 23 studies). Some studies (n = 5 studies) used a combination of VLE and WBE. The rest of the studies used WBE (n = 13 studies), CBE (n = 1 studies), multimodality (n = 9 studies), and mobile (n = 1 studies). Moreover, the studies

TABLE VI.   DISTRIBUTION OF PUBLISHED WORK PER SAMPLE
POPULATION

Sample population	Num. of studies
Middle school students	2
High school students	6
Undergraduate students	19
Graduate students	5
Educators (teachers, instructors, teaching assistants)	19

that were conducted within LMSs and who reported the type of the learning platform, have shown that Moodle was the most used type of a learning platform (n = 6 studies), followed by Blackboard (n = 2 studies), and other social learning platforms (n = 8 studies) such as EdX [64], Khan Academy [65], Coursera [66], Elgg [67], THEOL [63], ed-Venture [68], video-based social platform [69] and virtual world platform [70]. Furthermore, with regard to the setting of the learning environment, some studies were conducted in purely digital learning environments (n = 15 studies), some in blended learning environments (n = 11 studies), and some in face-to-face learning environments (n = 10 studies), where students collaborated face-to-face using digital tools [71], multi-surface tabletops [42], or used mobile devices as a learning mediator [72].

Another important information for the learning context was the pedagogical approach that has been used. Majority of the papers (n = 26) did not report the use of a specific pedagogical approach, but the results for those who did, are shown in Fig. 5.
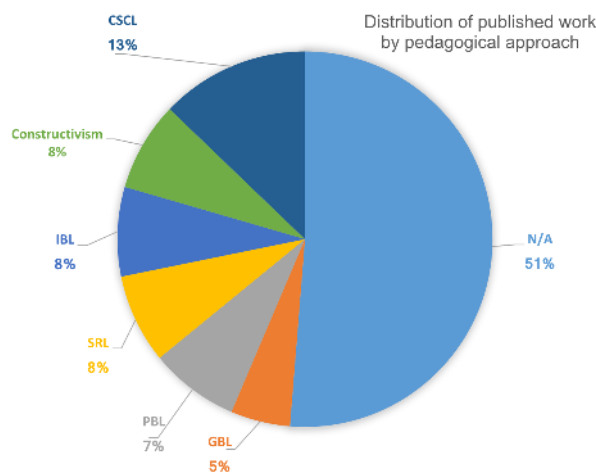

Fig. 5.   Distribution of published work by pedagogical approach

Within the setting of the learning environment, the authors also tried to systematically categorize the technology and tools used by the subjects during the studies. The most used technologies and tools are reported in Table VII.

Methodology and data analysis techniques. Regarding the type of methodology, the authors alluded to the type of methods used in the studies. Thus, majority of the studies used quantitative analysis (n = 21 studies), following mixed

TABLE VII.   TECHNOLOGY AND TOOLS USED IN THE PUBLISHED STUDIES

Technology and tools	Num. of studies
Software suits and web-based applications	8
Web 2.0 tools (wikis, chats, blogs, skype, social media tools, google, apps)	6
Dashboards and visualization tools	5
Kinetic sensors, EEG, eye-tracking devices	4
LA tools	3
Mobile phones/ iPad	2
Tabletops	2
Traditional tools (whiteboard, pen & paper, laptop, mouse, keyboard)	2

TABLE VIII.   DISTRIBUTION OF PUBLISHED WORK PER DATA COLLECTION METHODS

Data collection methods	Num. of studies
System logs (e.g. mobile logs, external 2.0 tools/apps)	19
LMS logs (e.g. course content, assessment, quiz, grades)	19
Surveys/Questionnaires	16
Documentation (e.g. notes, diary, peer evaluation, e-mail)	13
Interviews	10
Artifacts	7
Observations	6
Video/audio recordings	6
Discussion boards	5
Multimodal (e.g. EEG headset, kinetic sensor, eye–tracking)	4
Workshops/meetings	2
Data set originally collected for different research purpose	1

methods analysis (n = 14 studies), and qualitative analysis (n = 5 studies). The findings show that quantitative analysis is still the dominant methodology in learning analytics research. This shows that most of the analysis are done using learners' data from the LMSs or data gathered using any software suit or web-based application. Considering the data collection methods used in the studies, it can be concluded that various methods have been used and the most practiced data collection methods are presented in Table VIII.

With respect to data analysis, learning analytics adopts a wide range of techniques from statistics, data mining, text analysis, and social network analysis. For the purpose of this study, the authors decided to follow the general classification presented in [56] based on qualitative and quantitative methods used. Thus, the authors decided to report the results as a descriptive or inferential statistics, of which the inferential statistics can also be divided into parametric and non-parametric statistics. Consequently, one of the most used techniques is regression, either linear or multiple, bivariate analysis (i.e. correlation), and cluster analysis. Table IX displays the classification of the most used data analysis techniques in the selected papers.

TABLE IX.   DISTRIBUTION OF PUBLISHED WORK PER DATA ANALYSIS TECHNIQUE

Data analysis techniques	Num. of studies
Inferential statistics	24
Parametric statistics	16
Non-parametric statistics	4
Descriptive statistics	15
Content analysis	5
Discourse analysis	3
Thematic analysis	3
Semantic analysis	1
Dispositional analysis	1

Research objective, behavior and performance. The selected papers mostly focused on the following research objectives:

- *Design and management of learning scenarios/activities*: examples cover work that underlines design of learning activities or work that evaluates the effects of teaching practices, re-design of learning activities, or more efficient management of learning scenarios (n = 23 studies);
- *Student learning behavior and engagement:* examples include work that covers monitoring or evaluating student learning behavior/patterns and engagement within the learning environment (n = 20 studies);
- *Usefulness of LA tools:* e.g. what are the benefits of using LA tools or how users perceive LA tools (n = 12 studies);
- *Teacher's professional development:* examples include work that focuses on increasing teacher's awareness or approaches that improve teacher's skills by incorporating new teaching techniques (n = 7 studies);
- *Improved orchestration:* e.g. enactment-time; adaptation of available pedagogical and technological resources in classrooms, or learner's support in online/blended learning environments to help them achieve their intended learning goals (n = 7 studies);
- *Student's self-reflection/self-assessment* (n = 7 studies);
- *Predictive modelling* (n = 6 studies);
- *Collaboration and interaction* (n = 5 studies);
- *Student assessment* (n = 4 studies);
- *Overall user satisfaction* (n = 3 studies);
- *Student retention* (n = 2 studies);
- *Personalized learning* (n = 2 studies).

Moreover, some of the studies (n = 20 studies in total) reported the impact of learning analytics on subject's behavior (n = 18 studies) and the impact of learning analytics on learning performance (n = 9 studies). Appendix B lists the impact of learning analytics on subject's behavior and learning performance as reported by the authors of the selected studies. From the results it can be observed that learning analytics usage generally increased user awareness and user informedness in the learning environment [47], [67], [73], [74]. Next, usage of learning analytics assisted teachers to manage time better [47], [75], [42]; to identify problems in the course design [73], [72]; to follow student behavior and engagement with content overtime and apply informed changes to keep the level of instructional quality [76], [77]; to arrange the monitoring process according to their needs and think a priori about possible solutions [47], [69]; and to utilize real-time tracking and provide instantly corrective feedback [42], [68], [7], [71]. On the other hand, usage of learning analytics helped students to apply diagnostic assessment following their own performance [74], [72], [78], [49]; to initiate and steer conversation with their peers or the teacher [42]; to better navigate and utilize the course content [77]; to reflect, self-direct their progress, and make informed decisions how to continue reaching learning goals [7], [78], [71]. However, the selected studies did not directly considered nor clearly reported any measurement of learning gains, or any other learning-

related constructs. Finally, it was interesting to notice that students perceived learning analytics as "setting reasonable parameters of what is appropriate, what makes sense, and what you are trying to do" [48]. The complete list from the selected studies is presented in Appendix B.

To answer the second research question, the authors looked at what learning analytics were used throughout the studies to inform learning design decisions. As one can notice in Appendix C, most of the analytics used in the studies are extracted analytics, presented back to the learners, usually as visualizations or dashboards. These analytics were used to quickly locate elements in the learning design or user's performance that deviated from a defined threshold values [49], [77], [79], [80]. On the other hand, few studies reported use of embedded analytics, that are integrated into the learning environment and could be used for real-time feedback [7], [47], [67], [71]. Furthermore, findings show that learning analytics mostly have been used to explore the physical, digital, and human elements in the learning ecosystem [42], [81], for post-course reflection and recommendation of resources [67], for prediction [82], [83], [46], [66], as a tool [71], [77], [84], to detect weaknesses and strengths in the learning design [5], [85], [86], and to define scripts with monitoring information [47]. Another conclusion that can be draw from the selected studies is the common use of time-related learning analytics to evaluate the design of the learning activities and learners' online behavior [87], [71]; score and frequency learning analytics to evaluate performance, difficulty of the learning content and assessment strategies [63], [88]; and aggregated learning analytics to identify trends and behaviors [5], [89].

The selected studies can also be distinguished from the usage approach of learning analytics. Some of the studies focused on aggregating data from large data sets (usually blended courses or MOOCs) for the purpose of finding patterns within the data that can be applied in different contexts and among various modules [46], [5], [65], [90]. Moreover, these studies also aim to explore how learning design links to student online behavior [91]. Others, focused on actionable analytics in authentic learning settings [43], [42], [48], [49]. Both approaches are highly relevant as they supplement the research work to address the main challenge in learning analytics, i.e. to deliver actionable feedback, derived from theory, the learning context, and the methods in which the learning activities are situated [90], [5].

Moreover, the authors evaluated learning analytics indicators according to five perspectives: individual student, group, course, content, teacher; and six data sources: student generated data, local data, academic profile, evaluation, course-related performance, course meta-data. The findings revealed that most of the learning analytics are linked to individual students and their actions; students interaction with the learning content; or group's interactions; while learning analytics gathered from educators are less common. In regard to the data sources, majority of the studies used student generated data and evaluation data, neglecting course meta-data and local data. The complete list of the extracted learning analytics (152 in total) from the selected studies is presented in Appendix D.

Finally, to summarize, learning analytics can support dy-namic and data-driven learning design decisions if they are: collected from multiple data sources, modes, or learning settings; embedded in teachers' everyday practice; and a regular part of students' learning processes. Researchers that reported usage of multimodal learning analytics, [84], [42], [75], [92], [93] could support the analysis and study of complex, open-ended, hands-on learning experiences in authentic learning settings. This way, researchers have gained more insights into users' needs and expectations as a promising way to support dynamic and real-time learning design activities. Moreover, use of learning analytics only from digital footprints contextualized with qualitative data from users' experiences, and framed usage process during the learning activities, could also be a promising area that needs to be explored more [7], [47], [78]. A good example is Wise et al. [48] study that underline the importance of setting up the frame for learning analytics use, interpretation and decision-making as an integral part of students and teachers everyday activities tied to goals and expectations. Another example is Rodríguez-Triana et al. [47] study that focused on explicit guidance on how to use, interpret and reflect using learning analytics findings to adequately refine and re–design learning activities.

V. DISCUSSION AND FUTURE RESEARCH DIRECTIONS

Learning design as a field has produced methodologies, tools, and representations to assist educators in designing learning activities, while learning analytics holds the metrics, analysis, and reporting of data, to inform and influence the design process, and ensure appropriate refinement. Looking at the publication distribution per year, the interplay between learning analytics and learning design has gained expanding interest in the TEL community for further exploration of their alignment and conditional maturation.

In general, the rational behind the use of learning analytics for learning design is to discover learning phenomena (e.g. moment of learning or misconception) and design improved and pedagogically sound learning environments utilizing technology and resources. Thus, the majority of the studies focused on:

- utilization of learning analytics tools from which analytics were extracted and used to further develop the tools, as to offer better practical support and informed decision making [71], [77], [76], [94], [74], [93], [73], [80];
- development of frameworks that add a theoretical clarity to the learning process, identify analytics metrics, and create guidelines and recommendations that can inform the design of learning activities [43], [87], [75], [48], [69], [49], [81], [64].

A. Interpretation of the results with respect to the first research question

The summary from the selected studies have shown that students and teachers as users of learning analytics became more knowledgeable about their learning behaviors and progress. For example, they could anticipate how lack of information or emergence of problems could affect their activities and expectations [49], [67], [7], [88]. As a result, increased awareness

lead to improved and informed decision-making, and potential growth in users' skills and competencies [42], [47]. Next, although the usage of learning analytics could have impact on user's behavior, there are chances that users can utilize learning analytics to monitor their progress but not necessarily change their behavior [69]. However, learning analytics could also do the opposite, enact unintentional change [7]. Moreover, due to the lack of usage of qualitative studies, researchers might fail to gain awareness of the scale of learning analytics metrics that students find it useless or inaccurate [48], [7]. *Thus, the research community could benefit if there is a taxonomy of identified learning analytics that are likely to create an impact on user's behavior (positive or negative) and induce change in situ*. Also, the research community needs educators to take part in the process of design and implementation of learning analytics, as this could fill in the gap between what various learning analytics metrics present and what educators actually need [47], [87], [43].

Furthermore, from the analysis of the selected studies (see Appendix A), it can be noted that the use of technology (learning analytics tools, digital learning platforms, sensor-based tools) has increased the range for data collection. Researchers can collect data not just from the digital learning platforms or tools, but also from the physical spaces where learning is happening [42], [31], [93]. However, although data collection from physical spaces is becoming more common, the multimodal approaches to analyze the learning experiences in the physical spaces are not yet widespread [31]. In addition, the results from this review study support the findings from the orchestration research, that modelling and supporting teacher's orchestration in technology-rich physical and digital learning environments develops a great practical importance for the research community [31].

When it comes to methodology, the results show that quantitative studies still take precedence over mixed methods and qualitative studies due to the abundance of user activity data from LMSs. Therefore, the most practiced data collection methods were system logs, followed by surveys, while the most practiced data analysis techniques were derived from inferential statistics. However, simple clicking behavior in a LMS is a poor proxy for the actual learning behavior students have [82]. *This heavy reliance on digital footprints, often using a single platform as a source of data, focuses only on factors connected to numeric methods and hinders the holistic approach to understand the learning process as an ecosystem*.

In regard to the setting of the learning environment, it can be noted that most of the research is performed in supervised environments (i.e virtual learning environments/ learning management systems). This allows researchers to easily acquire data and answer questions related to quantitative measures of use or differentiation between learning offerings [38]. Hence, researchers use various indicators from system logs to understand:

- what type of content students use, how often, and how much time they spend interacting with the content? [90], [76], [63];
- what type of indicators help teachers to notice behaviors

and patterns? [65], [72], [71];
- what metrics and frequencies, such as visit duration and number of sessions are useful for teachers to reflect upon the analytics? [5], [83].

However, when it comes to more complex questions, such as "which effects do specific learning offerings have on collaborative learning processes?", researchers need more than just a quantitative data [78], [88]. These questions, which are concerned with user satisfaction, preferences, or needs, are partially answered, due to shortage of qualitative data. As a result, there is a misalignment between the information generated by learning analytics tools with the needs, problems, and concerns that teachers have regarding learning designs and learning activities [43], [95]. Likewise, students face the same problem as users of analytics, due to:

1) the lack of metrics which are pedagogically valuable [96];
2) failure of thoughtful design to encourage and shape analytics use [48];
3) failure to ground and tie analytics metrics to learners' goals and expectations [7].

These issues, place students as passive users of learning analytics, falling to empower them to take responsibility and regulate their own learning and performance. Considering the findings from the selected studies, the authors want to point out to:

- the use of broad set of complementary metrics (e.g. multimodal learning analytics) [42], [84], [75], [92] that will incorporate more aspects that characterize the learners and help researchers learn more about the learning process,
- and the importance of grounding learning analytics in theory before introducing and integrating the same into the learning environment [2], [94], [97].

Next, throughout the reviewed papers one can observe that there are few learning analytics designs that are grounded in explicit pedagogical models [47], [67], [7]. Many pedagogical models are implicit, or it happens the study to not even focus on any particular model. However, researchers need to explore and document various pedagogical factors that contribute to student success (e.g., learning gain and engagement during learning) so that subsequent work can have reference point from past research. For example, Berlands et al. [71] reported that using pedagogically defined learning analytics grounded in theory (i.e., Zone of Proximal Development specifically for learning Computer Science), provided a strong proof-of-concept that real-time support using personalized and theoretically grounded learning analytics can improve student performance, increase and maintain quality, and engage students to work together on more complex problems. *Consequently, in future, it is advised to consider the context that critically shapes learning [98], and investigate if and how data-driven design-decisions can be validated. In other words, the researchers need to find a way to explicitly label efficient learning design decisions build on particular data-driven analytics that are theoretically justified* [97], [2], [58].

B. Interpretation of the results with respect to the second research question

In the last ten years, learning is becoming more blended and distributed across different learning environments and contexts. Hence, turning to holistic approaches and considering how learning takes place today, becomes a necessity. Along these lines, the integration, aggregation, and harmonization of learning-related data across multiple sources and spaces has the potential to offer rich evidence-driven design that could amplify humans' learning capacities [42], [75]. Consequently, if learning analytics integration is neglected, future learning design will be guided from poor insights drawn from limited learning activities.

Learning analytics collected from the digital learning spaces are often complemented with data coming from student management systems [73], [79] or from self-reported variables utilizing surveys [88], [99], [5]. On one hand, the dynamic and data-driven learning design decisions would not be guided solely by digital footprints of learning systems and numeric methods [88], but will incorporate more aspects and metrics that holistically characterize learners, their needs, and their expectations. On the other hand, utilizing analytics coming from a single and many times limited learning environment, have no added value when rich and more representative datasets are available [82]. Consequently, the combination of learning analytics coming from several digital and sometimes even physical footprints (e.g. learning environments, self-reported information or through use of sensor-based tools) could improve the interpretation of the observed learning behavior and the patterns noticed within the learning environment. This way, educators could properly scaffold the design process through informed decisions utilizing awareness and reflection.

When it comes to learning design activities (see Appendix C), most of the studies included in the review have used assimilative, assessment, communication, finding and handling of information, and productive activities, as defined in the learning design taxonomy [34]. This is also supported by Rienties and Nguyen research studies on the impact of learning design in university settings [5], [46], [86], [85]. They reported that learning design has a strong influence on learner's satisfaction [5]. In particular, communication and interactive activities engage students to spend more time in VLEs compared to productive and experiential activities [86]. Although the design of learning activities depend on the module [85], educators design learning activities differently over the time line of the course and reduce the variety of learning design activities when they introduce assessment activities [46].

Although most of the studies follow the traditional paradigm in which the teacher is the main end-user of learning analytics, more and more studies are reporting results utilizing visualized analysis to increase awareness among students for self-monitoring and self-reflection [6], [7], [72]. The results presented in Appendix C, show increase in the usage of learning analytics metrics that suppose to reinforce students' self-reflection. The main idea behind self-regulated learning derives from "context-specific processes that are selectively used by students to succeed in school" [100], when numerous educational institutions have moved to "technology-rich" environment in which learning and teaching expands beyond the walls of university lecture settings. However, there is a limited research on how students accept, interpret, and use learning analytics to follow and improve their own performance [7], [6], [42].

Moreover, from Appendix D one can observe that only few analytics indicators have been explicitly devised to collect and present educator's data (i.e., data gathered from teachers, that allows teachers to reflect upon their teaching practices and course design) [47], [67], [63], [75], [69], [90], [42]. This shows that there is some educator's data stored in databases, mostly data from teacher's interactions with content or students, or log data from teacher's use of dashboards. However, the selected studies failed to report course meta-data, such as data regarding the course structure, resources or teaching concepts. *Consequently, lack of teacher's data and course meta-data, limit educator's opportunities to reflect on their teaching activities, pedagogical practices, the quality of the learning content and the interactions, that might lead to improvements in their professional development and dissemination of their hands-on experiences [101].*

What it is really interesting to be further applied at the intersection of learning analytics and learning design, is finding meaningful behavior patterns and interactions that can be "mapped back to the planned learning activities to explain why certain peaks and falls occur over time" [102], [83], [42]. For this, information regarding the course meta-data, the intended teaching concepts, and a feedback loop will be necessary. As Reinmann has noted, "more is needed than just data to "discover" meaningful relations" [2]. Thus, the research community could join efforts to develop conceptual framework that could model the complexities of the learning process towards comprehensible analytics and visualization requirements to transform the learning design into a teacher-led enquiry-based practice [43], [33]. Furthermore, what is often overlooked and underestimated but immensely important to educators, is the need for explicit guidance on how to use, interpret, and reflect on the learning analytics findings to adequately refine and re–design learning activities [47], [67], [101]. *A direction towards closing this gap is to consider establishing a participatory culture of design, and a habit among educators to see learning design as an inquiry process and learning analytics as a part of the teaching culture* [28], [33].

Finally, not all of the studies implicitly reported that they seek to visualize the impact of learning design activities to learners and educators from which they collected the data [90], [102], [63]. At present, the authors agree with Clow's argument that the learning analytics cycle can be completed even if interventions in learning design does not reach the learners from whom originally the data was generated as long as it is used to apply improvements for the next cohort [58]. However, for future work, researchers could try and close the cycle with the same group of learners due to authenticity and context. Returning the data gathered from authentic settings to students or teachers from which it has been collected, could assist in getting practical comments for further improvements

to local decisions in which the end-users were involved [47], [71], as well as to increase the awareness among the users for the applicability of learning analytics.

C. Theoretical and practical implications

One of the biggest challenges for researchers and practitioners will be to create a strong relation among learning analytics and learning theories to empower reflective practices [10], [1], [103]. Theory validates associations and patterns between digital traces and learning outcomes (i.e. learning-related constructs) that can trigger a learning process to reach specific goals [2]. Applying the consolidated model of theory, design and data science, proposed by [1], could bring invaluable insights to researchers. This means that researchers will know what data to collect in order to understand whether certain learning processes are activated, and what learning outcomes are associated with what design decisions [47]. Failing to consider context, could lead to misinterpretations of findings and limit the design replication in various learning settings [15], [98].

As it can be observed from the findings, much of the work in learning analytics has been related to development of visualizations [47], [72], [71], [75], [76], [42], [49], [74], [87]. However, there is a limited empirical evidence that visually presenting analyzed data could promote desirable learning practices and increased understanding in learning analytics interpretation [94]. Wise et al. [7] conducted a study grounded in computer supported collaborative learning (CSCL) theory, and proposed embedded and extracted learning analytics to generate visualizations for the purpose of self-directed and self-regulated learning. This work is particularly significant as it aims to establish feedback mechanism between students and teachers by fostering dialog where learning analytics is the conversation starter. In addition, Lockyer et al. [4] highlighted that learning design should consider the inclusion of learning analytics within a specific pedagogical design as a mean to encourage learning. Both studies underline the need for theoretical integration of learning analytics and learning design, that could be seen as a promising start to connect theory, design, and data to inform research and practice of learning analytics [1]. Also, this review study can serve as a springboard on furthering the alignment between learning analytics and learning design.

Another theoretical implication this study presents is the development of intermediate-level body of knowledge in learning analytics for learning design. Intermediate-level knowledge (some researchers also refer to it as strong concepts) includes "solution-oriented pieces of generative knowledge, residing on a level of abstraction between instances and theories" [104]. In the systematic literature review, we identified that the majority of the selected studies apply design-based research to inform practice and advance theory by iterations, as suggested by [2]. Thus, applying concept-driven approach to design-based research could potentially lead to construction of generative design knowledge [104]. Thinking in this direction, developing and applying strong concepts requires skills for thorough understanding of particular design scenarios and situations,

that goes beyond just mapping novel experiences from single studies. This should encourage researchers to engage in longitudinal studies that will not only change the way we think about studies in educational research (e.g. from short-term interventions to continuous monitoring) [47] but also whether and when certain scaffolds can gradually be removed to avoid cognitive load [105], how to design learning activities differently over time [86], [46], and expertise reversal effects [106]. Theories and principles are, by definition, formulated on a high level of abstraction, so that they can be applied to many different situations (generalization), which are then presented as instantiations of the abstract notions. To elicit principles and pave the way towards a unified theory of learning analytics for learning design, researchers need to triangulate research findings across the different case studies and meta-analyze the empirical knowledge. This will allow researchers to move from instances/case studies, to intermediate-level knowledge, and then to theory construction.

Finally, the authors would like to present a conceptual model towards learning analytics for learning design (LA4LD) taxonomy (see Figure 6). This taxonomy should derive classification from existing research and from the review study. Thus, on one hand, the proposed conceptual model incorporates the already existing learning design taxonomy proposed by [34] which identifies seven broad types of learning activities to guide design and creativity in the design process. On the other hand, the authors employed the Campbell and Oblinger's [107] five-step model of learning analytics: capture, report, predict, act, and refine and tried to map the findings from the review study in accordance with the model.

As described in the methodology section, subsection III.F, the authors proposed second level branches in the "taxonomy tree" that correspond to the five steps of learning analytics [107]. In particular, capture corresponds to the data collection; report corresponds to the techniques used to provide feedback using learning analytics; predict corresponds with the purpose to use learning analytics for prediction of grades or failures; act corresponds with the applied actions; and refine corresponds with interventions and re-design of learning scenarios, tools, etc. This is a proposition on how the findings from the selected studies can be used to derive classification and establish selection criteria for extracting pedagogically valuable learning analytics metrics from specific learning design activities. Furthermore, LA4LD taxonomy could summarize the objectives of the selected papers, as shown in the tree trunk (see Figure 6). These objectives are derived from the synergy between learning analytics and learning design, as reported in the selected studies. The proposed LA4LD taxonomy, offers a springboard for further work; so other conceptual knowledge development endeavors can utilize it. Thus, the proposed LA4LD taxonomy can be seen as an essential step towards future empirical research, and a support tool to researchers in establishing a connection between learning analytics for learning design and theory.
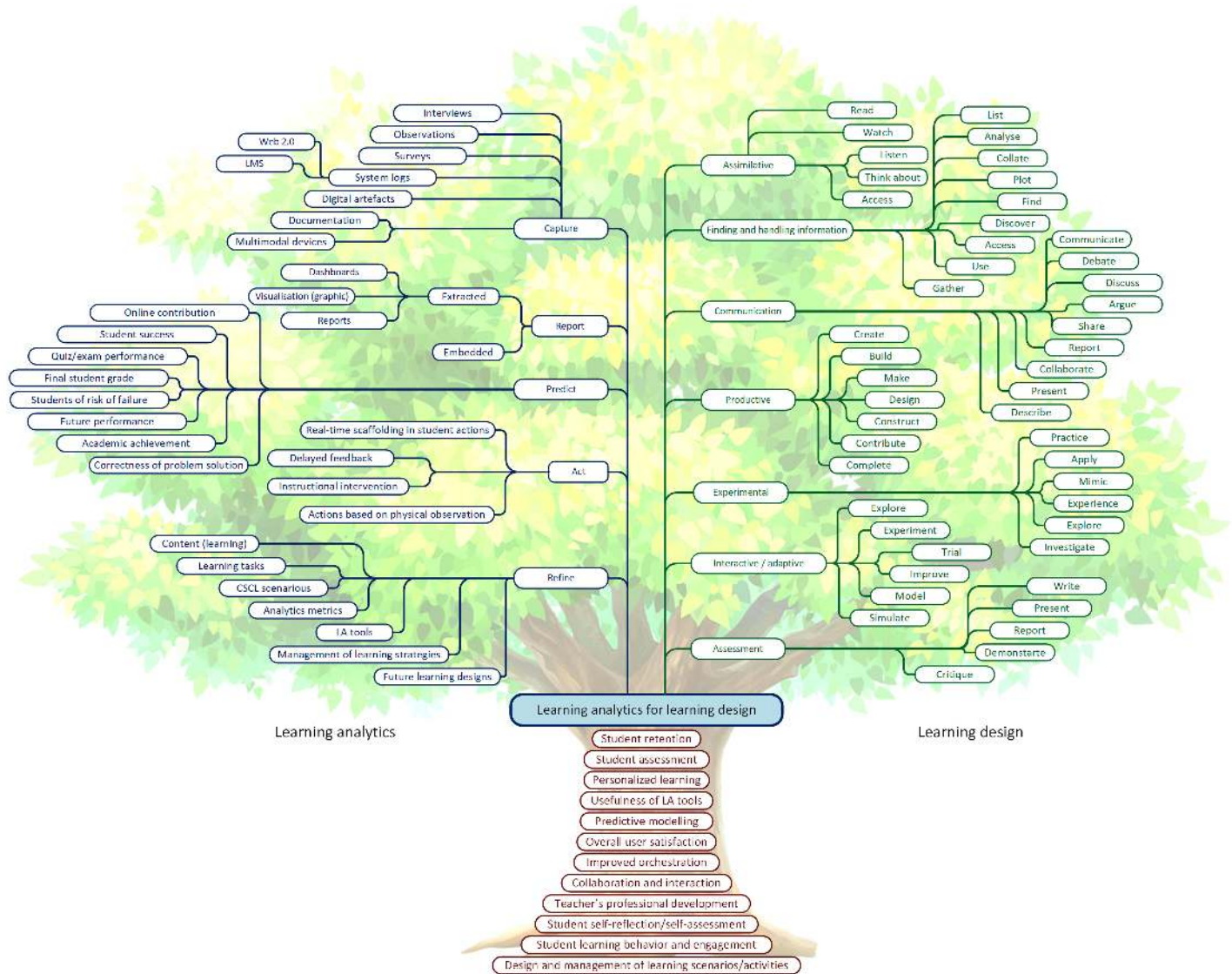
Fig. 6.   Learning analytics for learning design taxonomy

D. Future work

Based on the reviewed papers, the authors want to offer the following checklist for future work on learning analytics for learning design:

- provide details about the learning environment and the used pedagogical approaches, where improvements in learning design experiences based on learning analytics outcomes will be measured [47];
- indicate how learning analytics metrics offer insights into learning processes and can be theoretically grounded for meaningful interpretation to inform theory and design [6];
- evaluate and compare what learning design patterns and learning phenomena make learning effective [5];

- evaluate and denote student learning outcomes, or any other learning-related constructs [71];
- evaluate and denote the impact of learning analytics outcomes on learning design decisions and experiences [72];
- evaluate and denote how educators are planning, designing, implementing, and evaluating learning design decisions [101];
- provide common guidance on how to use, interpret and reflect on the learning analytics to adequately refine and redesign learning activities [7]

E. Limitations

The main limitations of this review can be seen as:

- bias in the selection of databases, journals, and publications due to possible subjectivity and lack of relevant information;
- bias in the search string because keywords are discipline and language-specific;
- shortage to draw more general conclusions since the focus was on empirical research;
- bias and inaccuracy in data extraction as it was performed only by the two authors;
- bias from interpretation of some findings, methods, or approaches, as some parts of the methodology from the selected studies were not described accurately.

However, the authors attempted to ensure unbiased review process by developing a research protocol in advance with already defined research questions. The search string was developed using the research questions and considering a possible lack of standardization in keywords as they can be discipline- and language-specific. Furthermore, the authors performed a search of state-of-the-art in TEL, in terms of journals, databases, and previous review studies. Finally, the selected methodology (i.e. Systematic Literature Review) is an in-depth study of a relatively narrow area using specific and pointed research questions that entails certain limitations. Other approaches like Systematic Mapping Study [108] might not go deep into an area, but create a map of a wide research field. Despite the limitations of the selected methodology, it is a well-accepted and widely used literature review method [9], [18] in TEL providing certain assurance of the results.

VI. CONCLUSION

The current review demonstrates the present landscape at the intersection between learning analytics and learning design. Learning analytics and learning design are two complementary fields within educational technology, that together hold the promise to optimize the learning process, and contribute to the creation of more meaningful learning experiences, tools, and evidence-based practices. The authors analyzed 43 peer-reviewed articles selected from the literature within the period of 2010-2017. This review aimed to explore what learning analytics have been used to inform learning design decisions, and what were the main design approaches researchers embraced over the last seven years. Furthermore, the authors explored the convergence of learning analytics and learning design within the reported learning settings, the pedagogical contexts, the data collection methods, the data analysis techniques, and the reported research objectives.

The review has shown that future research should consider developing a framework on how to capture and systematize learning design data, and follow what learning design choices made by educators influence subsequent learning activities and performances over time. Furthermore, it is of utmost importance to theoretically ground the learning analytics because: 1) the choice of methods and analysis used in the studies should be driven by theory and practice of learning analytics; 2) the findings from the studies should be used to inform theory and design. Addressing these elements could help in further maturation of the fields of learning analytics

and learning design, and provide foundation for longitudinal and comparative studies among various educational contexts. Furthermore, educators and researchers need to leverage the use of learning analytics and focus on developing students' skills and natural predispositions by designing personalized learning and feedback, while decreasing assimilative activities as traditional lecturing, reading, or watching videos. Future learning needs to be direct it towards personalizing learners' experiences and adapting it to their strengths, interest, and aspirations. Also, educators need to re-think their role of simply being providers of knowledge, to designers and facilitators of learning.

As a summary, the authors like to highlight the main idea of aligning learning analytics with learning design, as an essential condition to create more meaningful tools, methods, and representations of data for educators and learners. Thus, this alignment would lead to improved and informed learning decisions, and towards the development of design principles and knowledge between data representation and data-driven actions.

APPENDIX

- **Appendix A.** Results from the coding process
- **Appendix B.** Learning analytics impact on subject's behavior and learning performance
- **Appendix C.** Learning analytics for learning design
- **Appendix D.** Learning analytics indicators categorized according perspective and data sources

ACKNOWLEDGMENT

REFERENCES

[1] D. Gašević, V. Kovanović, and S. Joksimović, "Piecing the learning analytics puzzle: a consolidated model of a field of research and practice," *Learning: Research and Practice*, vol. 3, no. 1, pp. 63–78, 2017.

[2] P. Reimann, "Connecting learning analytics with learning research: the role of design-based research," *Learning: Research and Practice*, vol. 2, no. 2, pp. 130–142, 2016.

[3] Y. Mor and B. Craft, "Learning design: reflections upon the current landscape," *Research in learning technology*, vol. 20, no. sup1, p. 19196, 2012.

[4] L. Lockyer, E. Heathcote, and S. Dawson, "Informing pedagogical action: Aligning learning analytics with learning design," *American Behavioral Scientist*, vol. 57, no. 10, pp. 1439–1459, 2013.

[5] B. Rienties and L. Toetenel, "The impact of learning design on student behaviour, satisfaction and performance: A cross-institutional comparison across 151 modules," *Computers in Human Behavior*, vol. 60, pp. 333–341, 2016.

[6] D. Gašević, N. Mirriahi, S. Dawson, and S. Joksimović, "Effects of instructional conditions and experience on the adoption of a learning tool," *Computers in Human Behavior*, vol. 67, pp. 207–220, 2017.

[7] A. Wise, Y. Zhao, and S. Hausknecht, "Learning analytics for online discussions: Embedded and extracted approaches," *Journal of Learning Analytics*, vol. 1, no. 2, pp. 48–71, 2014.

[8] G. Siemens, "Learning analytics: The emergence of a discipline," *American Behavioral Scientist*, vol. 57, no. 10, pp. 1380–1400, 2013.

[9] Z. Papamitsiou and A. A. Economides, "Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence," *Journal of Educational Technology & Society*, vol. 17, no. 4, p. 49, 2014.

[10] R. Ferguson, "Learning analytics: drivers, developments and challenges," *International Journal of Technology Enhanced Learning*, vol. 4, no. 5-6, pp. 304–317, 2012.

[11] G. Siemens and P. Long, "Penetrating the fog: Analytics in learning and education." *EDUCAUSE review*, vol. 46, no. 5, p. 30, 2011.

[12] G. Siemens and D. Gašević, "Guest editorial-learning and knowledge analytics." *Educational Technology & Society*, vol. 15, no. 3, pp. 1–2, 2012.

[13] P. Long, G. Siemens, G. Conole, and D. Gašević, in *Proceedings of the 1st International Conference on Learning Analytics and Knowledge (LAK'11)*. ACM, 2011.

[14] K. E. Arnold and M. D. Pistilli, "Course signals at purdue: Using learning analytics to increase student success," in *Proceedings of the 2nd international conference on learning analytics and knowledge*. ACM, 2012, pp. 267–270.

[15] D. Gašević, S. Dawson, T. Rogers, and D. Gasevic, "Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success," *The Internet and Higher Education*, vol. 28, pp. 68–84, 2016.

[16] G. D. Magoulas, Y. Papanikolaou, and M. Grigoriadou, "Adaptive web-based learning: accommodating individual differences through system's adaptation," *British journal of educational technology*, vol. 34, no. 4, pp. 511–527, 2003.

[17] A. Klašnja-Milićević, B. Vesin, M. Ivanović, Z. Budimac, and L. C. Jain, *E-Learning Systems: Intelligent Techniques for Personalization*. Springer, 2016, vol. 112.

[18] Z. Papamitsiou and A. A. Economides, "Learning analytics for smart learning environments: a meta-analysis of empirical research results from 2009 to 2015," *Learning, Design, and Technology: An International Compendium of Theory, Research, Practice, and Policy*, pp. 1–23, 2016.

[19] M. Van Harmelen and D. Workman, "Analytics for learning and teaching," *CETIS Analytics Series*, vol. 1, no. 3, pp. 1–40, 2012.

[20] K. L. Gustafson and R. M. Branch, *Survey of instructional development models*. ERIC, 1997.

[21] R. A. Reiser, "A history of instructional design and technology: Part ii: A history of instructional design," *Educational technology research and development*, vol. 49, no. 2, pp. 57–67, 2001.

[22] P. M. Ghislandi and J. E. Raffaghelli, "Forward-oriented designing for learning as a means to achieve educational quality," *British Journal of Educational Technology*, vol. 46, no. 2, pp. 280–299, 2015.

[23] D. Laurillard, "Technology enhanced learning as a tool for pedagogical innovation," *Journal of Philosophy of Education*, vol. 42, no. 3-4, pp. 521–533, 2008.

[24] M. Maina, C. Brock, and M. Yishay, *The art & science of learning design*. Springer, 2015, no. 9.

[25] G. Conole, *Designing for learning in an open world*. Springer Science & Business Media, 2012, vol. 4.

[26] E. Dobozy, "Typologies of learning design and the introduction of a 'ld-type 2' case example," *eLearning Papers–special edition*, 2012.

[27] R. Koper, "Current research in learning design," *Educational Technology & Society*, vol. 9, no. 1, pp. 13–22, 2006.

[28] D. Persico and F. Pozzi, "Informing learning design with learning analytics to improve teacher inquiry," *British Journal of Educational Technology*, vol. 46, no. 2, pp. 230–248, 2015.

[29] L. P. Prieto, Y. Dimitriadis, C.-K. Looi *et al.*, "Orchestration in learning technology research: evaluation of a conceptual framework," 2015.

[30] P. Dillenbourg, "Design for classroom orchestration," *Computers & Education*, vol. 69, pp. 485–492, 2013.

[31] L. Prieto, K. Sharma, Ł. Kidzinski, M. Rodríguez-Triana, and P. Dillenbourg, "Multimodal teaching analytics: Automated extraction of orchestration graphs from wearable sensor data," *Journal of computer assisted learning*, vol. 34, no. 2, pp. 193–203, 2018.

[32] L. P. Prieto Santos *et al.*, "Supporting orchestration of blended cscl scenarios in distributed learning environments," 2012.

[33] L. Prieto, M. Rodríguez-Triana, R. Martínez-Maldonado, Y. Dimitriadis, and D. Gašević. (2018) Orchestrating learning analytics (orla): supporting the adoption of learning analytics at the practitioner level. [Online]. Available: osf.io/y2p7j

[34] B. Rienties, Q. Nguyen, W. Holmes, and K. Reedy, "A review of ten years of implementation and research in aligning learning design with learning analytics at the open university uk," *Interaction Design and Architecture (s)*, vol. 33, pp. 134–154, 2017.

[35] K. Verbert, N. Manouselis, H. Drachsler, and E. Duval, "Dataset-driven research to support learning and knowledge analytics," 2012.

[36] D. Gašević, S. Dawson, and G. Siemens, "Let's not forget: Learning analytics are about learning," *TechTrends*, vol. 59, no. 1, pp. 64–71, 2015.

[37] P. Goodyear and Y. Dimitriadis, "In medias res: reframing design for learning," *Research in Learning Technology*, vol. 21, no. 1, p. 19909, 2013.

[38] A. L. Dyckhoff, "Implications for learning analytics tools: A meta-analysis of applied research questions," *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 3, no. 1, pp. 594–601, 2011.

[39] S. Joksimović, A. Manataki, D. Gašević, S. Dawson, V. Kovanović, and I. F. De Kereki, "Translating network position into performance: importance of centrality in different network configurations," in *Proceedings of the sixth international conference on learning analytics & knowledge*. ACM, 2016, pp. 314–323.

[40] M. A. Chatti, A. L. Dyckhoff, U. Schroeder, and H. Thüs, "A reference model for learning analytics," *International Journal of Technology Enhanced Learning*, vol. 4, no. 5-6, pp. 318–331, 2012.

[41] W. Greller and H. Drachsler, "Translating learning into numbers: A generic framework for learning analytics," *Journal of Educational Technology & Society*, vol. 15, no. 3, p. 42, 2012.

[42] R. Martinez-Maldonado, B. Schneider, S. Charleer, S. B. Shum, J. Klerkx, and E. Duval, "Interactive surfaces and learning analytics: data, orchestration aspects, pedagogical uses and challenges," in *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. ACM, 2016, pp. 124–133.

[43] A. Bakharia, L. Corrin, P. de Barba, G. Kennedy, D. Gašević, R. Mulder, D. Williams, S. Dawson, and L. Lockyer, "A conceptual framework linking learning design with learning analytics," in *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. ACM, 2016, pp. 329–338.

[44] J. Dunlosky, K. A. Rawson, E. J. Marsh, M. J. Nathan, and D. T. Willingham, "Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology," *Psychological Science in the Public Interest*, vol. 14, no. 1, pp. 4–58, 2013.

[45] Y. Mor, R. Ferguson, and B. Wasson, "Learning design, teacher inquiry into student learning and learning analytics: A call for action," *British Journal of Educational Technology*, vol. 46, no. 2, pp. 221–229, 2015.

[46] Q. Nguyen, B. Rienties, L. Toetenel, R. Ferguson, and D. Whitelock, "Examining the designs of computer-based assessment and its impact on student engagement, satisfaction, and pass rates," *Computers in Human Behavior*, vol. 76, pp. 703–714, 2017.

[47] M. J. Rodríguez-Triana, A. Martínez-Monés, J. I. Asensio-Pérez, and Y. Dimitriadis, "Scripting and monitoring meet each other: Aligning learning analytics and learning design to support teachers in orchestrat-

ing cscl situations," *British Journal of Educational Technology*, vol. 46, no. 2, pp. 330–343, 2015.

[48] A. F. Wise, J. M. Vytasek, S. Hausknecht, and Y. Zhao, "Developing learning analytics design knowledge in the "middle space": The student tuning model and align design framework for learning analytics use," *Online Learning*, vol. 20, no. 2, 2016.

[49] E. Koh, A. Shibani, J. P.-L. Tan, and H. Hong, "A pedagogical framework for learning analytics in collaborative inquiry tasks: an example from a teamwork competency awareness program," in *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. ACM, 2016, pp. 74–83.

[50] L. Lockyer and S. Dawson, "Where learning analytics meets learning design," in *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. ACM, 2012, pp. 14–15.

[51] S. Keele *et al.*, "Guidelines for performing systematic literature reviews in software engineering," in *Technical report, Ver. 2.3 EBSE Technical Report. EBSE*. sn, 2007.

[52] T. Dybå and T. Dingsøyr, "Empirical studies of agile software development: A systematic review," *Information and software technology*, vol. 50, no. 9-10, pp. 833–859, 2008.

[53] T. Greenhalgh, *How to read a paper*. BMJ Publishing Group London, 2001, vol. 2.

[54] B. A. Kitchenham, S. L. Pfleeger, L. M. Pickard, P. W. Jones, D. C. Hoaglin, K. El Emam, and J. Rosenberg, "Preliminary guidelines for empirical research in software engineering," *IEEE Transactions on software engineering*, vol. 28, no. 8, pp. 721–734, 2002.

[55] L. P. Prieto, M. Holenko Dlab, I. Gutiérrez, M. Abdulwahed, and W. Balid, "Orchestrating technology enhanced learning: a literature review and a conceptual framework," *International Journal of Technology Enhanced Learning*, vol. 3, no. 6, pp. 583–598, 2011.

[56] A. Bhattacherjee, *Social science research: Principles, methods, and practices*. Global Text Project, 2012.

[57] J. P. Campbell, P. B. DeBlois, and D. G. Oblinger, "Academic analytics: A new tool for a new era," *EDUCAUSE review*, vol. 42, no. 4, p. 40, 2007.

[58] D. Clow, "The learning analytics cycle: closing the loop effectively," in *Proceedings of the 2nd international conference on learning analytics and knowledge*. ACM, 2012, pp. 134–138.

[59] O. D. Space. Glossary of teaching approaches. [Online]. Available: http://portal.opendiscoveryspace.eu/tr-activity/22-macibampieeja-vardnica-669764

[60] A. L. Dyckhoff, V. Lukarov, A. Muslim, M. A. Chatti, and U. Schroeder, "Supporting action research with learning analytics," in *Proceedings of the Third International Conference on Learning Analytics and Knowledge*. ACM, 2013, pp. 220–229.

[61] R. E. Clark, "Constructing a taxonomy of media attributes for research purposes," *AV Communication Review*, vol. 23, no. 2, pp. 197–215, 1975.

[62] D. Gašević, D. Djuric, and V. Devedžic, *Model driven architecture and ontology development*. Springer Science & Business Media, 2006.

[63] J. Ma, X. Han, J. Yang, and J. Cheng, "Examining the necessary condition for engagement in an online learning environment based on learning analytics approach: The role of the instructor," *The Internet and Higher Education*, vol. 24, pp. 26–34, 2015.

[64] C.-U. Lei, X. Hou, T. T. Kwok, T. S. Chan, J. Lee, E. Oh, D. Gonda, Y.-C. A. Yeung, and C. Lai, "Advancing mooc and spoc development via a learner decision journey analytic framework," in *Teaching, Assessment, and Learning for Engineering (TALE), 2015 IEEE International Conference on*. IEEE, 2015, pp. 149–156.

[65] J. A. Ruipérez-Valiente, P. J. Muñoz-Merino, D. Leony, and C. D. Kloos, "Alas-ka: A learning analytics extension for better understanding the learning process in the khan academy platform," *Computers in Human Behavior*, vol. 47, pp. 139–148, 2015.

[66] B. K. Pursel, L. Zhang, K. W. Jablokow, G. Choi, and D. Velegol, "Understanding mooc students: motivations and behaviours indicative

of mooc completion," *Journal of Computer Assisted Learning*, vol. 32, no. 3, pp. 202–217, 2016.

[67] P. A. Haya, O. Daems, N. Malzahn, J. Castellanos, and H. U. Hoppe, "Analysing content and patterns of interaction for improving the learning design of networked learning environments," *British Journal of Educational Technology*, vol. 46, no. 2, pp. 300–316, 2015.

[68] Á. Serrano-Laguna, J. Torrente, P. Moreno-Ger, and B. Fernández-Manjón, "Application of learning analytics in educational videogames," *Entertainment Computing*, vol. 5, no. 4, pp. 313–322, 2014.

[69] J. Castellanos, P. Haya, and J. Urquiza-Fuentes, "A novel group engagement score for virtual learning environments," *IEEE Transactions on Learning Technologies*, 2016.

[70] R. Baker, J. Clarke-Midura, and J. Ocumpaugh, "Towards general models of effective science inquiry in virtual performance assessments," *Journal of Computer Assisted Learning*, vol. 32, no. 3, pp. 267–280, 2016.

[71] M. Berland, D. Davis, and C. P. Smith, "Amoeba: Designing for collaboration in computer science classrooms through live learning analytics," *International Journal of Computer-Supported Collaborative Learning*, vol. 10, no. 4, pp. 425–447, 2015.

[72] J. Melero, D. Hernández-Leo, J. Sun, P. Santos, and J. Blat, "How was the activity? a visualization support for a case of location-based learning design," *British Journal of Educational Technology*, vol. 46, no. 2, pp. 317–329, 2015.

[73] B. Florian-Gaviria, C. Glahn, and R. F. Gesa, "A software suite for efficient use of the european qualifications framework in online and blended courses," *IEEE Transactions on Learning Technologies*, vol. 6, no. 3, pp. 283–296, 2013.

[74] K. Kitto, M. Lupton, K. Davis, and Z. Waters, "Designing for student-facing learning analytics," *Australasian Journal of Educational Technology*, vol. 33, no. 5, pp. 152–168, 2017.

[75] R. Martinez-Maldonado, K. Yacef, and J. Kay, "Tscl: A conceptual model to inform understanding of collaborative learning processes at interactive tabletops," *International Journal of Human-Computer Studies*, vol. 83, pp. 62–82, 2015.

[76] A. Tervakari, K. Kuosa, J. Koro, J. Paukkeri, and M. Kailanto, "Teachers' learning analytics tools in a social media enhanced learning environment," in *Interactive Collaborative Learning (ICL), 2014 International Conference on*. IEEE, 2014, pp. 355–360.

[77] M. A. Chatti, M. Marinov, O. Sabov, R. Laksono, Z. Sofyan, A. M. F. Yousef, and U. Schroeder, "Video annotation and analytics in coursemapper," *Smart Learning Environments*, vol. 3, no. 1, pp. 1–21, 2016.

[78] A. Pardo, R. A. Ellis, and R. A. Calvo, "Combining observational and experiential data to inform the redesign of learning activities," in *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*. ACM, 2015, pp. 305–309.

[79] N. Li, V. Marsh, and B. Rienties, "Modelling and managing learner satisfaction: Use of learner feedback to enhance blended and online learning experience," *Decision Sciences Journal of Innovative Education*, vol. 14, no. 2, pp. 216–242, 2016.

[80] F. Martin and A. Ndoye, "Using learning analytics to assess student learning in online courses," *Journal of University Teaching & Learning Practice*, vol. 13, no. 3, p. 7, 2016.

[81] K. Thompson, D. Ashe, L. Carvalho, P. Goodyear, N. Kelly, and M. Parisio, "Processing and visualizing data in complex learning environments," *American Behavioral Scientist*, vol. 57, no. 10, pp. 1401–1420, 2013.

[82] D. T. Tempelaar, B. Rienties, and B. Giesbers, "Verifying the stability and sensitivity of learning analytics based prediction models: An extended case study," in *International Conference on Computer Supported Education*. Springer, 2015, pp. 256–273.

[83] N. Z. Zacharis, "A multivariate approach to predicting student out-

comes in web-enabled blended learning courses," *The Internet and Higher Education*, vol. 27, pp. 44–53, 2015.

[84] K. Thompson, "Using micro-patterns of speech to predict the correctness of answers to mathematics problems: an exercise in multimodal learning analytics," in *Proceedings of the 15th ACM on International conference on multimodal interaction*. ACM, 2013, pp. 591–598.

[85] Q. Nguyen, B. Rienties, and L. Toetenel, "Mixing and matching learning design and learning analytics," in *International Conference on Learning and Collaboration Technologies*. Springer, 2017, pp. 302–316.

[86] ——, "Unravelling the dynamics of instructional practice: a longitudinal study on learning design and vle activities," 2017.

[87] F. Martin, A. Ndoye, and P. Wilkins, "Using learning analytics to enhance student learning in online courses based on quality matters standards," *Journal of Educational Technology Systems*, vol. 45, no. 2, pp. 165–187, 2016.

[88] A. Pardo, F. Han, and R. A. Ellis, "Combining university student self-regulated learning indicators and engagement with online learning events to predict academic performance," *IEEE Transactions on Learning Technologies*, vol. 10, no. 1, pp. 82–92, 2017.

[89] R. F. Kizilcec, C. Piech, and E. Schneider, "Deconstructing disengagement: analyzing learner subpopulations in massive open online courses," in *Proceedings of the third international conference on learning analytics and knowledge*. ACM, 2013, pp. 170–179.

[90] S. Joksimović, D. Gašević, T. M. Loughin, V. Kovanović, and M. Hatala, "Learning at distance: Effects of interaction traces on academic achievement," *Computers & Education*, vol. 87, pp. 204–217, 2015.

[91] L. Corrin, P. G. de Barba, and A. Bakharia, "Using learning analytics to explore help-seeking learner profiles in moocs," in *Proceedings of the seventh international learning analytics & knowledge conference*. ACM, 2017, pp. 424–428.

[92] M. Worsley and P. Blikstein, "Leveraging multimodal learning analytics to differentiate student learning strategies," in *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*. ACM, 2015, pp. 360–367.

[93] K. Pantazos and R. Vatrapu, "Enhancing the professional vision of teachers: A physiological study of teaching analytics dashboards of students' repertory grid exercises in business education," in *System Sciences (HICSS), 2016 49th Hawaii International Conference on*. IEEE, 2016, pp. 41–50.

[94] A. F. Wise, "Designing pedagogical interventions to support student use of learning analytics," in *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*. ACM, 2014, pp. 203–211.

[95] K. Avramides, J. Hunter, M. Oliver, and R. Luckin, "A method for teacher inquiry in cross-curricular projects: Lessons from a case study," *British Journal of Educational Technology*, vol. 46, no. 2, pp. 249–264, 2015.

[96] S. B. Shum and R. Ferguson, "Social learning analytics," *Journal of educational technology & society*, vol. 15, no. 3, p. 3, 2012.

[97] M. Neelen and P. Kirschner. Where are the learning sciences in learning analytics research? [Online]. Available: https://3starlearningexperiences.wordpress.com/2017/10/17/where-are-the-learning-sciences-in-learning-analytics-research/

[98] P. H. Winne, "How software technologies can improve research on learning and bolster school reform," *Educational Psychologist*, vol. 41, no. 1, pp. 5–17, 2006.

[99] S. Freitas, D. Gibson, C. Du Plessis, P. Halloran, E. Williams, M. Ambrose, I. Dunwell, and S. Arnab, "Foundations of dynamic learning analytics: Using university student data to increase retention," *British Journal of Educational Technology*, vol. 46, no. 6, pp. 1175–1188, 2015.

[100] B. J. Zimmerman, "Academic studing and the development of personal skill: A self-regulatory perspective," *Educational psychologist*, vol. 33, no. 2-3, pp. 73–86, 1998.

[101] S. McKenney and Y. Mor, "Supporting teachers in data-informed educational design," *British journal of educational technology*, vol. 46, no. 2, pp. 265–279, 2015.

[102] B. Rienties, L. Toetenel, and A. Bryan, "Scaling up learning design: impact of learning design activities on lms behavior and performance," in *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*. ACM, 2015, pp. 315–319.

[103] A. F. Wise and D. W. Shaffer, "Why theory matters more than ever in the age of big data." *Journal of Learning Analytics*, vol. 2, no. 2, pp. 5–13, 2015.

[104] K. Höök and J. Löwgren, "Strong concepts: Intermediate-level knowledge in interaction design research," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 19, no. 3, p. 23, 2012.

[105] J. S. Brown, A. Collins, and P. Duguid, "Situated cognition and the culture of learning," *Educational researcher*, vol. 18, no. 1, pp. 32–42, 1989.

[106] S. Kalyuga, "Expertise reversal effect and its implications for learner-tailored instruction," *Educational Psychology Review*, vol. 19, no. 4, pp. 509–539, 2007.

[107] D. Oblinger and J. Campbell, "Academic analytics, educause white paper," *Retrieved October*, vol. 20, p. 2011, 2007.

[108] B. A. Kitchenham, D. Budgen, and O. P. Brereton, "Using mapping studies as the basis for further research–a participant-observer case study," *Information and Software Technology*, vol. 53, no. 6, pp. 638–651, 2011.

Katerina Mangaroska is a PhD student at the Department of Computer Science at the Norwegian University of Science and Technology. Her primary research area is learning analytics. Currently she is engaged in the "Future Learning: Orchestrating 21st Century Learning Ecosystem using Analytics" project that aims to develop new knowledge how analytics allow us to better orchestrate different learning tools and practises, and design optimal learning environments. She is also working with eye-tracking to gain insights about user behavior in programming environments (e.g. how user processes information when learning to program or interacts with visual information during coding). Her other research interests center around learning design, intelligent tutoring systems, learning infrastructure for computing education, and human-computer interaction. Mangaroska is a Fulbright scholar.

Michail Giannakos is an associate professor of interaction design and learning technology at the Department of Computer Science, Norwegian University of Science and Technology and research director of the Centre for Excellent IT Education. His research interests center on making sense of user experiences and practices in order to redesign and optimize the education settings and systems. Giannakos is member on the executive board of IEEE Technical Committee on Learning Technology (IEEE TCLT). He has worked at several research projects funded by diverse sources like EC, Microsoft Research, Norwegian Research Council (NRC), NSF, DAAD and Cheng Endowment; Giannakos is also a recipient of a Marie Curie fellowship, the Norwegian CAREER award and he is one of the twenty outstanding academic fellows of NTNU.