Learning analytics for learning design: Towards evidence-driven decisions to enhance learning

Katerina Mangaroska and Michail Giannakos

Norwegian University of Science and Technology (NTNU), Trondheim, Norway
{katerina.mangaroska, michailg}@ntnu.no

Abstract. As the fields of learning analytics and learning design mature, the convergence and synergies between them become an important area for research. Collecting and combining learning analytics coming from different channels can clearly provide valuable information in designing learning. Hence, this paper intends to summarize the main outcomes of a systematic literature review of empirical evidence on learning analytics for learning design. The search was performed in seven academic databases, resulting in 38 papers included in the main analysis. The review demonstrates ongoing design patterns and learning phenomena that improve learning, by providing more comprehensive background of the current landscape of learning analytics for learning design and its impact on the current status of learning technologies. Consequently, future research should consider how to capture and systematize learning design data. Moreover, it should evaluate and document what learning design choices made by educators using what learning analytics techniques influence learning experiences and learning performances over time.

Keywords: Learning Analytics, Learning Design, Empirical Studies.

1 Introduction

Due to the pervasion of learning technologies, the use of learning analytics to discover important learning phenomena (e.g., moment of learning or misconception) and portray learners' experiences and behaviors, becomes evident and commonly accepted. At present, there are various analytic methods and e-learning tools that can be used to improve the learning experience [8]. However, without contextual interpretation of the data collected with e-learning tools, learning analytics capabilities are limited. From this perspective, learning design is utterly important as it provides the framework for analyzing and interpreting learner's behavior and data. Learning design defines the educational objectives and the pedagogical approaches that educators can reflect upon, take decisions and make improvements. Moreover, learning design "document the sequence of learning tasks, the resources, and the sequence of teaching methods" as main premises for reusability and transferability of good practices across educational contexts [5]. Yet, past research was focused on "conceptualizing learning design principles, without evaluating what happens after the design process" [9]. In addition, several studies have tried to understand and improve the learning design experiences by utilizing learning analytics, but only few of them tried empirically to show that learning analytics and learning design are complementary research areas that together have significant impact on the learning process [8]. Consequently, a research work is missing to measure what learning design decisions affect learning behavior and stimulate productive learning environment.

To bridge the gap, this paper centers in a systematic literature review with aim to examine the intersection between learning analytics and learning design, and provide important insights beyond the specific research findings within the individual discipline. 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 explore the extent to which learning analytics can support dynamic and data-driven learning design decisions.

2 Methodology

To answer the research questions, the authors decided to follow the guidelines for systematic literature review in software engineering [4]. In fact, before conducting the review, the authors developed a review protocol to reduce researcher bias and to keep a clear scope of the study. The search was performed in iterations in five main academic electronic databases in Technology Enhanced Learning (TEL): ACM DL, IEEE Xplore, SpringerLink, Science Direct, and Wiley, and two additional databases, SAGE and ERIC. The second cycle, included an independent search in the top ten educational and technology journals listed in the Google metrics sub-category: Educational Technology. The third and final cycle included search in the reference section for each selected paper in order to find additional relevant papers (i.e. the snowball technique). For the search, a research string was generated using combination of three terms: "*analytics*" AND "*design*" AND "*learning*". The literature search was performed from mid-October 2016 till mid-December 2016.

The process of evaluation and selection of papers consisted of several stages and followed inclusion/exclusion criteria defined by the authors. As a result, 288 papers were selected for the second stage of the systematic review. After the second stage and following the CAPS checklist, a total of 38 papers were retrieved, read it entirely, coded, and critically assessed by two researchers (see the list of papers here: https://figshare.com/articles/Bibliography_for_systematic_review_on_learning_analytics_and_learning_design_docx/4871690).

3 Findings

Regarding RQ1, the authors established the following parameters:

Sample population. The predominant sample population consists of undergraduates (n = 14) and educators (n = 14) including teachers and instructors, followed by high school (n = 6), graduates (n = 5), and middle school (n = 3) students.

Learning setting. Majority of the studies (n = 16) were conducted within VLEs and/or LMSs, followed by web-based environments (n = 6), MOOCs (n = 4), and multimodality (n = 4), computer-based environments (n = 3), video-based environment (n = 2), cognitive tutors (n = 1), and mobile learning environment (n = 1). The selected studies were conducted in purely digital (n = 19) and blended (n = 11) learning settings.

Learning scenarios. Formal learning was addresses in most of the studies (n = 27) rather than informal and non-formal learning.

Pedagogical approach. Majority of the papers (n = 19) did not refer to a specific pedagogical approach, but those that reported, includes: problem-based learning (n = 4), project-based learning (n = 4), game-based learning (n = 3), and CSCL (n = 3).

Technology and tools. Most used are applications specifically developed to test types of learning analytics, and web 2.0 social media tools (e.g. wiki, chat, google apps). Some studies reported use of devices for multimodal human-computer interaction such as tabletops, kinetic sensors, EEG, and eye tracking.

Type of methodology. Majority of the studies used quantitative analysis (n = 23), mixed methods (n = 8), and qualitative analysis (n = 7).

Data collection methods. Most practiced data collection methods are user activity LMS logs (n = 14) or logs coming from web 2.0 social media tools (n = 14), followed by analytics coming from questionnaires (n = 12), interviews (n = 6), observations (n = 4), and multimodal analytics (n = 3).

Data analysis techniques. Most popular techniques used are inferential statistics (n = 20) especially regression and clustering, followed by descriptive statistics (n = 9), content analysis (n = 6), and correlation (n = 5). Other used techniques included data mining (n = 3), social network analysis (n = 3), discourse analysis (n = 3), thematic (n = 2), and text analysis (n = 2). Reported only once are grounded theory, phenomenology, semantic analysis, sentiment analysis, and heuristic mining.

Research objectives. Student learning behavior (n = 8), collaboration and interaction (n = 7), and student assessment (n = 7) are the primary research objectives. Next are design and management of learning scenarios (n = 5), student retention (n = 5), learning performance (n = 5 studies), predictive modelling (n = 5), and student monitoring and engagement (n = 5).

Regarding RQ2, the authors analyzed the learning analytics used in the studies, and the learning design challenges/decisions that have been proposed.

One of the most known practices in learning analytics is collecting and analyzing historical and current LMS user activity data to study students' learning paths (i.e. trajectories) [3]. Another type of analytics often used in the selected studies are ready to be visualized learning analytics that emphasize informed and real-time feedback [6]. Next, the authors observed analytics coming from student's digital artifacts, such as artifacts from project-work or video-based learning settings. Furthermore, a very interesting finding was the expanded use of combined learning analytics coming from different data sources [14]. Finally, few studies reported the importance and use of multi-modal sensor data like kinetics, EEG, eye-movement, speech and body-movement [7, 13].

When it comes to learning design, one of the main challenges found in the analyzed studies is the seamless integration of design tools and strategies for lessons planning with tools for monitoring and analysis [10]. Moreover, the results also demonstrated the need for seamless integration of multimodal data (e.g. integrating physiological measures for better understanding learners' actions and experience) and the need to differentiate what can truly be designed in the learning environment [7, 13]. However, the current landscape of learning design depicts the visual representations of the outcomes from learning analytic, especially dashboards, as an easy to understand and concise way of presenting valuable information [11, 1].

4 Discussion and Conclusion

From the results described above, we can extract several main findings and propose directions for future research studies. The results show that quantitative methodology still takes precedence over mixed methods and qualitative methodology due to the abundance of user activity data from LMSs. However, simple clicking behavior in a LMS is a poor proxy for the actual learning behavior students have [12]. *This heavy reliance on log analysis, often using a single platform as a source of data is one of the primary issues that needs to be address in near future.* Furthermore, learning is becoming more blended and distributed across different learning environments and contexts, making it impossible to holistically understand the process of learning if integration is neglected. Therefore, the authors want to highlight *the importance of learning analytics integration and aggregation of learning-related data across multiple sources for designing informed and optimal learning strategies*.

Although most of the studies follow the traditional paradigm in which the teacher is the main user monitoring students, more and more studies are reporting results from using visualization analytics to increase awareness among students for self-monitoring and self-reflection [2]. The main idea is to help learners improve self–diagnostic of their own performance and seek solutions accordingly. In fact, for this issue there is a limited research on how students interpret and use learning analytics to follow their own learning performance. Another important finding is that there is no accepted framework or agreed method for research which learning analytics are used for what learning design challenges, nor examples of sharing and reuse of current methods and practices across various educational contexts. *This is one of the hallmarks of young fields, as well as the lack of longitudinal and comparative studies*.

On the other side, one of the most striking findings of this review is the lack of studies which directly consider and measure student learning gains, or any other learning-related constructs. Another unexpected finding is the shortage of studies on how educators are planning, designing, implementing, and evaluating learning design decisions [9]. Furthermore, there is an insufficient number of studies that consider using learning analytics to intentionally design learning activities that support collaboration and cooperation among students, rather than just following learner performance over time [3]. Finally, 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. *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* [8].

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;
- evaluate and compare what learning design patterns and learning phenomena make learning effective;
- evaluate and denote student learning gains, or any other learning-related constructs;
- evaluate and denote the impact of learning analytics outcomes on learning design decisions and experiences;
- evaluate and denote how educators are planning, designing, implementing, and evaluating learning design decisions;
- provide common guidance on how to use, interpret and reflect on the learning analytics to adequately refine and redesign learning activities.

This 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. 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. Finally, educators and researchers need to leverage the use of learning analytics and focus on developing students' skills and natural predispositions by designing personalized feedback and tailored learning while decreasing assimilative activities as traditional lecturing, reading or watching videos.

Acknowledgments. This work was supported by the Research Council of Norway under the project FUTURE LEARNING (255129/H20).

References

1. Berland M., Davis D., Smith C. P.: AMOEBA: Designing for collaboration in computer science classrooms through live learning analytics. International Journal of Computer-Supported Collaborative Learning 10 (4), 425-447 (2015).

2. Gašević D., Mirriahi N., Dawson S., Joksimović S.: Effects of instructional conditions and experience on the adoption of a learning tool. Computers in Human Behavior 67, 207-220 (2017).
3. Joksimović S., Gašević D., Loughin T. M., Kovanović V., Hatala M.: Learning at distance: Effects of interaction traces on academic achievement. Computers & Education 87, 204-217 (2015).
4. Kitchenham B., Charters S. (2007) Guidelines for performing systematic literature reviews in software engineering. EBSE Technical Report. , Ver. 2.3 edn. Keele University, United Kingdom
5. Lockyer L., Heathcote E., Dawson S.: Informing pedagogical action: Aligning learning analytics with learning design. American Behavioral Scientist 57 (10), 1439-1459 (2013).
6. Melero J., Hernández-Leo D., Sun J., Santos P., Blat J.: How was the activity? A visualization support for a case of location-based learning design. British Journal of Educational Technology 46 (2), 317-329 (2015).
7. Pantazos K., Vatrapu R.: Enhancing the professional vision of teachers: A physiological study of teaching analytics dashboards of students' repertory grid exercises in business education. In: 49th Hawaii International Conference on System Sciences (HICSS) 2016. IEEE, pp. 41-50
8. Persico D., Pozzi F.: Informing learning design with learning analytics to improve teacher inquiry. British Journal of Educational Technology 46 (2), 230-248 (2015).
9. Rienties B., Toetenel L.: The impact of learning design on student behaviour, satisfaction and performance: A cross-institutional comparison across 151 modules. Computers in Human Behavior 60, 331-341 (2016).
10. Rodríguez-Triana M. J., Martínez-Monés A., Asensio-Pérez J. I., Dimitriadis Y.: Scripting and monitoring meet each other: Aligning learning analytics and learning design to support teachers in orchestrating CSCL situations. British Journal of Educational Technology 46 (2), 330-343 (2015).
11. Schwendimann B., Rodriguez-Triana M., Vozniuk A., Prieto L., Boroujeni M., Holzer A., Gillet D., Dillenbourg P.: Perceiving learning at a glance: A systematic literature review of learning dashboard research. IEEE Transactions on Learning Technologies, (2016).
12. Tempelaar D., Rienties B., Giesbers B.: Verifying the stability and sensitivity of learning analytics based prediction models: An extended case study. In: International Conference on Computer Supported Education, 2015. Springer International Publishing, pp. 256-273
13. Worsley M., Blikstein P.: Leveraging multimodal learning analytics to differentiate student learning strategies. In: 5th International Conference on Learning Analytics and Knowledge, 2015. ACM, pp. 360-367
14. Zacharis N. Z.: A multivariate approach to predicting student outcomes in web-enabled blended learning courses. The Internet and Higher Education 27, 44-53 (2015).