The Features Learning Analytics Students Want the Most: Help Them Learn Over All Else

Mohammed Saqr[1], Sonsoles López-Pernas[1]

[1] *University of Eastern Finland, Yliopistokatu 2, 80100 Joensuu, Finland*

Abstract

To be effective, support based on learning analytics (LA) necessitates that students' attitudes, needs, and expectations are taken into account. Recently, research exploring students' needs and expectations has attracted the attention of LA researchers and practitioners driven by increasing focus on personalized learning and focus on the delivery of effective LA insights. Yet, most of such research comes from students who have a faint idea of LA, who do not firmly understand the potentials and the possible drawbacks inherent in LA. This current study aimed to fill this gap by surveying well-informed students —who completed an advanced course on LA— about the features they need from LA themselves. We also complemented our analysis with a network approach to understand the association and interplay between different needs. Our findings have shown that most of the students want LA features that help them perform their academic tasks: recommendations, feedback and reminders of deadlines. Students were most skeptical about comparing them with other students and suggesting other students as partners in academic work. The network analysis has confirmed such features and pointed out that resources and recommendations are the most central features that make students interested in LA. In a nutshell, students want LA to help them learn and support their learning journey over all else.

Keywords

learning analytics, expectations, survey, students

1. Introduction

Learning analytics (LA) is an interdisciplinary field that emerged more than a decade ago to take advantage of the increasingly digitized learning and modern data analysis methods. The most common definition of LA states that it "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" [1]. Ever since the field has started, there has been a growing array of data sources, methods, applications, and practices across diverse domains of research and practice [2]. Such diversity of LA as a field has enriched our understanding of students' learning and gave rise to several ways that we can help and support students in their learning. LA-based support necessitates that students' attitudes, needs, and expectations are taken into account [3]. Recently, research exploring students' needs and expectations has attracted the attention of LA researchers and practitioners driven by increasing focus on personalized learning, human-centered LA, and focus on the delivery of effective LA insights [3–5].

Most of such research comes from students who have a faint idea of LA and do not firmly understand the potentials and the possible drawbacks inherent in LA and datafication of learning [6]. The current study aims to fill this gap by surveying students who completed an advanced course on LA about the features they need from LA themselves. We believe that such students can offer a much-needed

Proceedings for the 14th International Conference on e-Learning 2023, September 28-29, 2023, Belgrade, Serbia
EMAIL: mohammed.saqr@uef.fi (A. 1); sonsoles.lopez@uef.fi (A. 2)
ORCID: 0000-0001-5881-3109 (A. 1); 0000-0002-9621-1392 (A. 2)

perspective from well-informed participants [3]. We augment our analysis with a network approach to study the interplay between the interest in several features and compute the most central features that students want.

2. Background

2.1. LA as means for supporting students

Whereas research in LA has grown almost exponentially over the past years, it has lagged behind in practice. The vast majority of LA research centers around research scenarios, hypothetical possibilities, and potential for support rather than offering support or actual intervention [2,7]. Nevertheless, some examples exist that have implemented LA based intervention, e.g., offered data driven feedback or LA dashboards to students. LA Dashboards (LADs) are probably the most common methods that have been used in LA in real-life practice [8,9]. LADs are graphical instruments designed to display LA insights to the students, teachers, or administrators. Such insights are assumed to support the decision-making process, augment cognition and raise awareness. For students —the subjects of this study— dashboards offer access to summaries of activities, comparative charts, and prescriptions. The premise is that the dashboards may help, e.g., raise students' awareness of their activities, help students reflect on their performance or help students make sense of the data presented to them [8]. Whereas research examining dashboards exists, the value and worth of LAD effectiveness is far from proven [8,9]. More importantly, what students want, or need and the LAD items that actually work is still an open question.

Students' support has also been offered through automated feedback on learning, assignments or learning activities [7,10,11]. Other examples of support may include personalized recommendation of learning resources, recommendations of future studies, learning strategies or learning activities. In other cases, LA-based feedback may ask students to contact their academic advisor or consult their teachers [12,13].

As we currently stand, several methods for offering automated support and feedback exit. Nonetheless, the efficacy of such intervention on students' learning outcomes —while promising— has not been proven beyond doubt. Furthermore, we know that certain features of LA support are more desirable by students than others and that students expect their universities and schools to offer some features more than others [3,14]. Yet, our knowledge comes mostly from students who have not tried LA or students who have only been told in a few lines what LA is [15]. It is fair to say here that students' opinions based on incomplete information can hardly reflect their true wants and needs. Therefore, our study aims to explore students' opinions about LA about the features they want from LA. Given the complexity and interdependence of features over each other, we complement this analysis by network analysis.

2.1.1. The network approach

Using a network approach to represent and analyze psychological and behavioral processes is consistent with the conceptualization of such phenomena as an ensemble of interacting components that interact and influence each other or what is known as a complex system [16,17]. Complex systems — as a theoretical grounding— align well with the fact that human psychological processes do not exist in vacuum isolated from each other but interact continuously together to result in emergent structures (e.g., learning strategy or attitude) [18,19]. Recently, methods for studying complex systems have advanced our understanding of several psychological phenomena, e.g., engagement, self-regulation and academic performance [20,21].

In particular, studying the interactions between variables has become possible with a subset of networks where the nodes represent the variables, and the edges are the statistical relationships; such networks are known as Pairwise Markov Random Fields (PMRF) [20,22]. PMRF networks have recently gained large-scale adoption under several names e.g., multivariate networks, psychological networks and Graphical Gaussian Models (GGMs) [23]. The methods have allowed researchers to study the structural relationships between variables in several contexts, for example, climate change, gene interactions, interdependence of behaviors or relationships between psychological phenomena [20]. We take advantage of network methods, namely, psychological networks to study the interplay between features and their centralities.

3. Methods

3.1.1. Data collection

The context of this study is a LA course at the University of Eastern Finland. The curriculum includes the principles of LA including the types of LA data, methods, learning theories and ethical concerns. The biggest focus of the course is placed on LA methods such as sequence analysis, process mining, social network analysis and predictive analytics. Students' put these methods to practice throughout a series of assignments in which they have to analyze real-life datasets using accessible tools (i.e., no coding skills are required). In addition, students learn about learning theories and ethical aspects of LA. The course has several assignments where students reflect on the literal, analyze datasets as well as offer a critique of the methods, and ethical approach of some papers. For more information about the context please see [15].

After the completion of the course, all students were surveyed about the features they want from LA to be offered to them. For that purpose, we used the validated Learning Analytics Features (LAF) questionnaire by Schumacher and Ifenthaler [3]. The questionnaire contains 15 items —corresponding to 15 LA features— that students had to agree or disagree with according to whether they needed the features in terms of learning, privacy concerns, and acceptance. A total of 114 students participated in the study throughout two editions of the course (2019-2020).

3.1.2. Data analysis

Students' responses to the questionnaire were analyzed using descriptive statistics, in order to understand general trends in responses. A Likert plot was also created to represent response distribution with the percentages of each

To understand the interplay and the associations between different features, we used psychological networks estimated using the methods described in [22,23]. Psychological networks are undirected PMRF networks where the nodes are variables and the edges are partial correlations. The presence of an edge indicates a partial correlation i.e., the two variables are conditionally dependent on each other after controlling for all other variables in the network (similar to regression) [24]. Moving forward, to avoid repetition, we will use the term "correlated" to indicate regularized partial correlation. The absence of a connection between two variables indicates that the two variables are independent from each other after controlling for all other variables in the network [22–24]. The network in our study was estimated according to the methods details in [22], In brief, a regularized partial correlation network was estimated from the survey variables using the R package [25]. The expected influence centrality was also computed to show the variables that have the most influence on the network connectivity or affect other variables' strength of interactions [24]. Given the similarity between survey items and the high correlation between several variables, we had to combine closely related constructs to balance sparsity over redundancy. Therefore, we combined time spent and work time load as time, learning resources recommendations, learning resources rating, extra learning resources, and revision learning resources as Resources. Recommend was made up of the personalized recommendations and suggestions of peers to work with. Status represents the timeline of status and goals. Planning includes deadline reminders and term scheduling. Support includes assignment feedback and prompts for self-reflection. Updates encompasses the newsfeed. Lastly, Compare is about comparison with other students.

3.2. Results

3.2.1. Questionnaire results

Table 1 shows the mean, median, and standard deviation of the responses to each item in the questionnaire. The distribution of responses can be seen in Figure 1 ordered by share of positive responses.

Table 1: Descriptive statistics of the questionnaire responses

Item	N	M	MED	SD
Comparison with fellow students	113	2.63	3	1.27
Considering the student's personal calendar for appropriate learning recommendations	113	3.56	4	1.08
Feedback for assignments	113	4.30	4	0.89
Learning recommendations for successful course completion	113	4.24	4	0.77
Further learning recommendations	112	4.00	4	0.95
Newsfeed	113	3.24	3	1.09
Prompts for self-assessment	113	3.85	4	0.93
Rating scales for learning material	113	3.96	4	0.82
Reminder for deadlines	113	4.24	5	1.02
Revision of previous learning content	111	3.77	4	0.85
Suggestion of learning partners	113	3.43	4	1.01
Term schedule	113	3.84	4	1.03
Time expected to complete a task or read a text	112	3.85	4	1.05
Time spent online	114	3.67	4	1.05
Timeline showing current status and goal	113	4.14	4	0.91

The first item of the questionnaire dealt with students' being aware of their academic progress and success in comparison with their fellow students. This was the most negatively scored item in the questionnaire, with a mean score of 2.63 (MED = 3, SD = 1.27), and close to half of the responses being negative. Given that the mean value was below the threshold of 3, it indicates that —on average— students do not want to be offered comparisons with other students.

When asked if they would like to obtain personalized learning recommendations based on their calendar commitments, 60% of the students replied affirmatively —with a 4 or a 5— (M = 3.56, MED = 4, SD = 1.08). The next item, which dealt with obtaining feedback for students' assignments, was the one with the highest mean score (M = 4.30, MED = 4, SD = 0.89) and had 90% of positive responses. This was closely followed by obtaining recommendations for successful course completion which, although it had a slightly lower mean score (M = 4.24, MED = 4, SD = 0.77), it had the highest share of positive responses (92%). Students were also interested —although not that much— in recommendations for further learning (M = 4.00, MED = 4, SD = 0.95), with 79% of positive responses. Having a newsfeed was not very attractive to students (M = 3.24, MED = 3, SD = 1.09), but also not detrimental, having the largest number of neutral responses (27%). Somewhat more interesting to students was having prompts for self-assessment (M = 3.85, MED = 4, SD = 0.93), with 70% of positive responses.

The availability of rating scales for the learning material was also quite well-received by the students (M = 3.96, MED = 4, SD = 0.82), with 78% of positive responses. A highly welcomed feature was having reminders for assignment deadlines (M = 4.24, MED = 5, SD = 1.02), with 84% of positive answers. Having opportunities for revising past content was moderately interesting to students (M = 3.77, MED = 4, SD = 0.85), with 71% of positive responses. Somewhat lower was the rating of the item about suggesting learning partners (M = 3.43, MED = 4, SD = 1.01), with only 58% positive responses. Having a scheduler for the school term was slightly higher (M = 3.84, MED = 4, SD = 1.03), with 68% positive responses. With an almost identical score, students were somewhat positive about knowing about the time expected to complete a task or read a text (M = 3.85, MED = 4, SD = 1.05, 69% positive responses). They were moderately interested in being aware of their time spent online (M = 3.67, MED = 4, SD = 1.05), with 69% positive responses. However, they were more interested in having a timeline showing their current status and goal (M = 4.14, MED = 4, SD = 0.91).

In summary, features that offer feedback, recommendations, or reminders were the most desired. Comparison with others and suggestions of other peers were the least desirable features

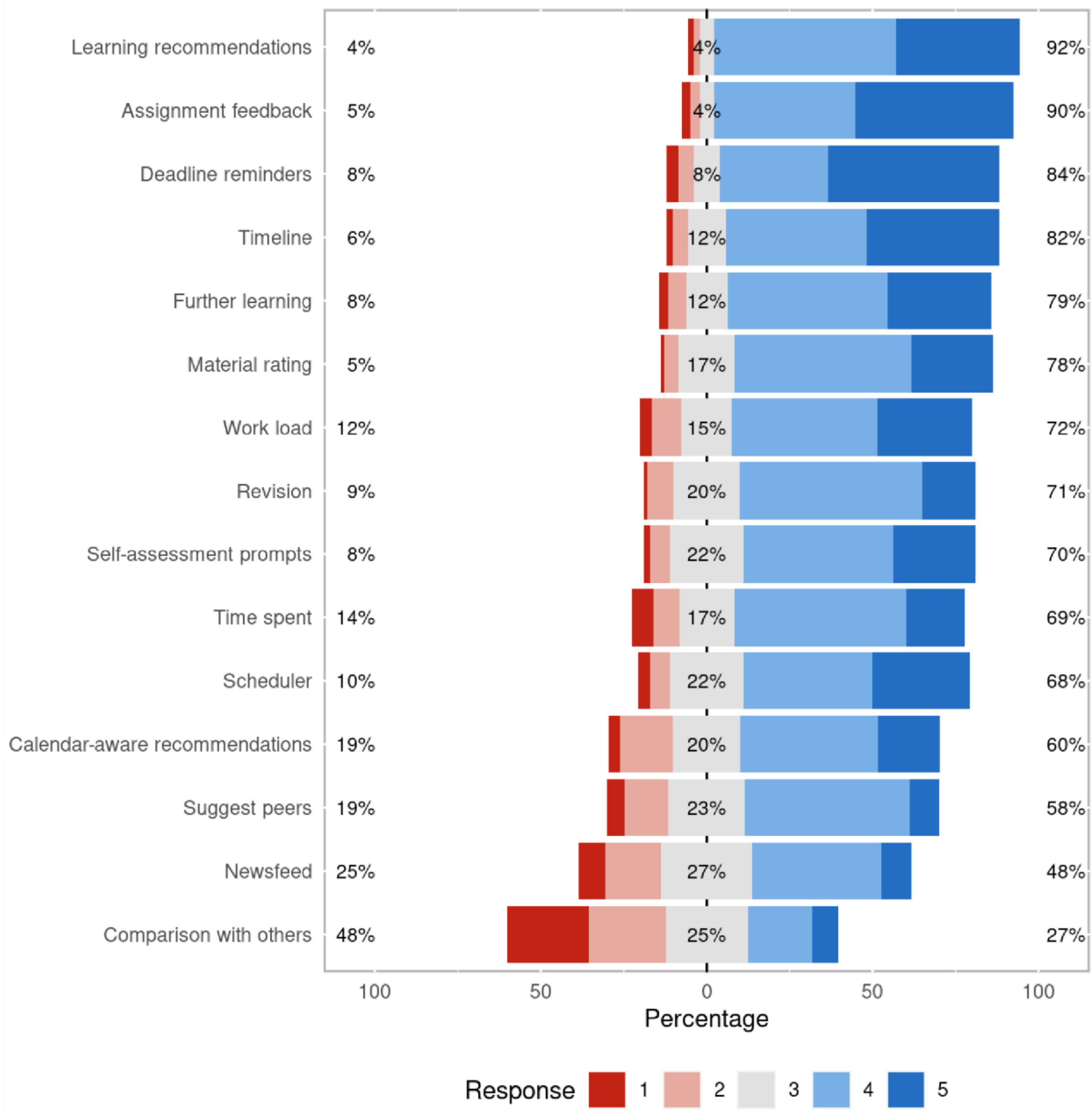

Figure 1: Likert plot of the survey responses. The percentages on the left represent the share of negative responses. (1, 2), the percentages in the middle represent neutral responses (3), and the percentages on the right side represent positive responses (4, 5)

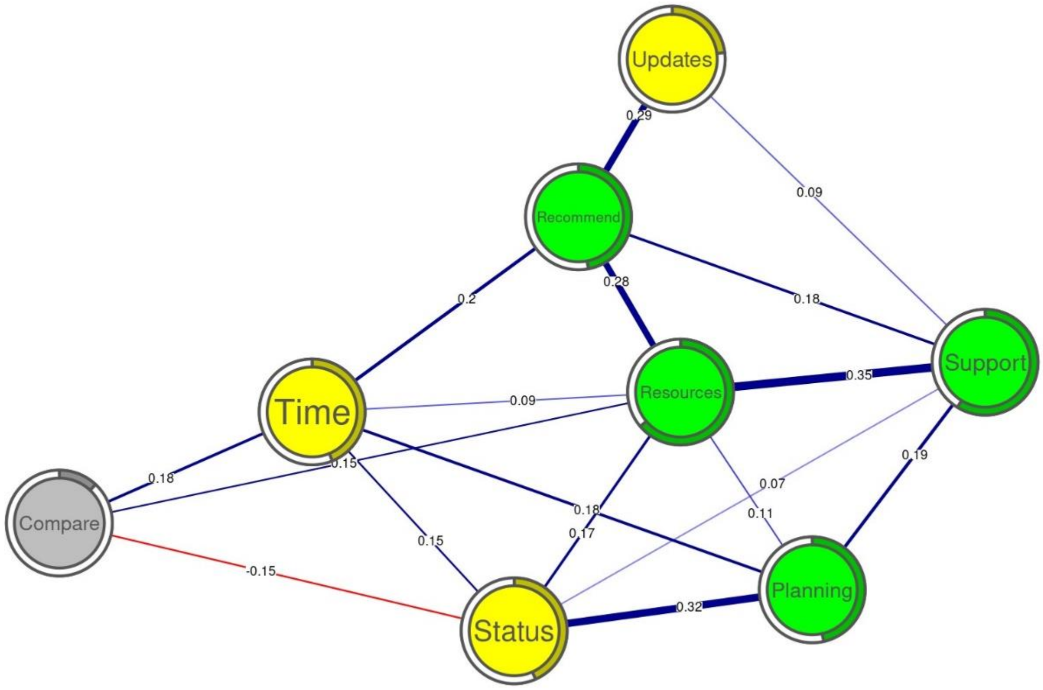

Figure 2: Network of associations between students' questionnaire responses. The thickness of the link is proportional to the partial correlation between the items. The color indicates the sign of the correlation: blue is positive and red is negative. The learning support nodes were colored in green, the nodes for tracking were colored yellow and comparison was colored grey

The partial correlation network (Figure 2) shows the features that are strongly connected and therefore interdependent on each other. We see features related to planning (scheduling and reminders) are strongly associated with tracking and time forming a strongly connected clique which can reflect the task planning, enactment and monitoring or regulation clique. Another strongly connected clique is formed by learning resources, recommendation and support which reflect the learning support clique. Updates or news feeds are only strongly connected to recommendations. Most importantly here is that comparison with others is negatively correlated with status, which means students want to know their own progress and at the same time and in the opposite direction do not want to be compared to other students. This finding emphasizes that students want LA but only features that help them learn not to put them in a race or competition with others.

The pie around each node shows the node's explain ability or how the connections of the node explain it. We see that learning resources and support are the most explainable nodes whereas compare is least explainable. This means that students' reluctance to be compared to other students is largely independent from their wish or need for other features of LA.

Figure 3: Expected influence centrality for each of the measured constructs

The expected influence centrality measures plot shows the features that drive the interest in other features from LA. As we can see, resources and recommendations are the most central features that have the highest expected influence on other features, followed by help planning and support. Interestingly, tracking, updates, and comparisons are the least central features that are less likely to kindle students' interest in LA.

4. Discussion

This study was performed to explore the opinions of LA students about LA features they may want. Our points of departure were that well-informed students can offer a more nuanced representation of students' attitude. We also complemented our analysis with a network approach to understand the association and interplay between different needs.

Our findings have shown that most of the students want LA features that help them do their academic work: recommendations, feedback and reminders of deadlines. Students were less enthusiastic about features that regulate or help them regulate their time such as time spent online. Most importantly, students were most skeptical about comparing them with other students and suggesting other students as partners for work. In fact, the score of comparison was below 3 which is consistent with a disagree statement. The network analysis has confirmed such features and pointed out that resources and recommendations are the most central features that make students interested in LA. In a nutshell, students want LA to help them learn and support their learning journey over all else.

A comparison with the previous research of Schumacher and Ifenthaler [3] shows that on all items, LA students want more of every feature of LA except for three features (course updates, prompts and revision). The idea that well-informed students want more indicates that when we educate students about LA, they are likely to want to use it in their classes. In particular, our students have rated the workload estimation and rating of learning resources far more favorably than Schumacher and Ifenthaler [3]. One explanation of this is that rating and load estimation have become more popular features of today's culture.

Table 2: A comparison between the mean of each score for the present study (M_{our}) and that of Schumacher and Ifenthaler [3] ($M_{original}$) and the difference ΔM as well as the effect size Cohen's d

Item	M_{our}	$M_{original}$	ΔM	Cohen's d
Comparison with fellow students	2.63	2.49	0.14	0.15
Personal calendar-based learning recommendations	3.56	3.50	0.06	0.07
Feedback for assignments	4.30	4.07	0.23	0.31
Learning recommendations for course completion	4.24	3.98	0.26	0.41
Further learning recommendations	4.00	3.69	0.32	0.42
Newsfeed	3.24	3.42	-0.18	-0.20
Prompts for self-assessment	3.85	4.07	-0.22	-0.33
Rating scales for learning material	3.96	3.31	0.65	0.83
Reminder for deadlines	4.24	4.20	0.04	0.06
Revision of previous learning content	3.77	4.12	-0.35	-0.56
Suggestion of learning partners	3.43	3.30	0.13	0.17
Term schedule	3.84	3.67	0.17	0.20
Time expected to complete a task or read a text	3.85	2.32	1.53	1.78
Time spent online	3.67	3.13	0.55	0.69
Timeline showing current status and goal	4.14	3.78	0.36	0.49

5. References

[1] G. Conole, D. Gašević, P. Long, G. Siemens, Message from the LAK 2011 general & program chairs, in: G. Conole, D. Gašević (Eds.), Proceedings of the 1st International Conference on Learning Analytics and Knowledge, Association for Computing Machinery (ACM), Banff, Canada, 2011.

[2] S. López-Pernas, K. Misiejuk, R. Kaliisa, M.Á. Conde, M. Saqr, Capturing the Wealth and Diversity of Learning Processes with Learning Analytics Methods, in: M. Saqr, S. López-Pernas (Eds.), Learning Analytics Methods and Tutorials: A Practical Guide Using R, Springer, 2024: p. in–press.

[3] C. Schumacher, D. Ifenthaler, C. Schumacher, D. Ifenthaler, Features students really expect from learning analytics, Comput. Human Behav. 78 (2018) 397–407.

[4] J. Merikko, K. Ng, M. Saqr, P. Ihantola, To Opt in or to Opt Out? Predicting Student Preference for Learning Analytics-Based Formative Feedback, IEEE Access. 10 (2022) 99195–99204.

[5] O. Viberg, L. Engström, M. Saqr, S. Hrastinski, Exploring students' expectations of learning analytics: A person-centered approach, Educ. Inf. Technol. 27 (2022) 8561–8581.

[6] M. Saqr, Big data and the emerging ethical challenges, Int. J. Health Sci. . 11 (2017) 1–2.

[7] S. Heikkinen, M. Saqr, J. Malmberg, M. Tedre, Supporting self-regulated learning with learning analytics interventions – a systematic literature review, Educ. Inf. Technol. (2022). https://doi.org/10.1007/s10639-022-11281-4.

[8] W. Matcha, N.A. Uzir, D. Gašević, A. Pardo, A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective, IEEE Trans. Learn. Technol. 13 (2020) 226–245.

[9] I. Jivet, M. Scheffel, H. Drachsler, M. Specht, Awareness Is Not Enough: Pitfalls of Learning Analytics Dashboards in the Educational Practice, in: Data Driven Approaches in Digital Education, Springer International Publishing, 2017: pp. 82–96.

[10] B.T.-M. Wong, K.C. Li, Learning Analytics Intervention: A Review of Case Studies, in: 2018 International Symposium on Educational Technology (ISET), IEEE, 2018: pp. 178–182.

[11] L. Lim, M. Bannert, J. van der Graaf, S. Singh, Y. Fan, S. Surendrannair, M. Rakovic, I. Molenaar, J. Moore, D. Gašević, Effects of real-time analytics-based personalized scaffolds on students' self-regulated learning, Comput. Human Behav. 139 (2023) 107547.
[12] B.T.-M. Wong, K.C. Li, A review of learning analytics intervention in higher education (2011–2018), Journal of Computers in Education. 7 (2020) 7–28.
[13] M. Saqr, U. Fors, M. Tedre, J. Nouri, How social network analysis can be used to monitor online collaborative learning and guide an informed intervention, PLoS One. 13 (2018) 1–22.
[14] A. Silvola, P. Näykki, A. Kaveri, H. Muukkonen, Expectations for supporting student engagement with learning analytics: An academic path perspective, Comput. Educ. 168 (2021) 104192.
[15] M. Saqr, S. López-Pernas, Why learning and teaching learning analytics is hard: An experience from a real-life LA course using LA methods, in: Proceedings of the Eleventh International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'23), Springer, 2024: p. in press.
[16] J. Malmberg, M. Saqr, H. Järvenoja, E. Haataja, H.J. Pijeira-Díaz, S. Järvelä, Modeling the Complex Interplay Between Monitoring Events for Regulated Learning with Psychological Networks, in: M. Giannakos, D. Spikol, D. Di Mitri, K. Sharma, X. Ochoa, R. Hammad (Eds.), The Multimodal Learning Analytics Handbook, Springer International Publishing, Cham, 2022: pp. 79–104.
[17] M. Saqr, M.J. Schreuder, S. López-Pernas, Why educational research needs a complex system revolution that embraces individual differences, heterogeneity, and uncertainty, in: M. Saqr, S. López-Pernas (Eds.), Learning Analytics Methods and Tutorials: A Practical Guide Using R, Springer, 2024: p. in–press.
[18] J.C. Hilpert, G.C. Marchand, Complex Systems Research in Educational Psychology: Aligning Theory and Method, Educ. Psychol. 53 (2018) 185–202.
[19] M. Koopmans, Education is a Complex Dynamical System: Challenges for Research, J. Exp. Educ. 88 (2020) 358–374.
[20] M. Saqr, Group-level analysis of engagement poorly reflects individual students' processes: Why we need idiographic learning analytics, Comput. Human Behav. (2023) 107991.
[21] J. Malmberg, M. Saqr, H. Järvenoja, S. Järvelä, How the monitoring events of individual students are associated with phases of regulation: A network analysis approach, Journal of Learning Analytics. 9 (2022) 77–92.
[22] S. Epskamp, L.J. Waldorp, R. Mõttus, D. Borsboom, The Gaussian Graphical Model in Cross-Sectional and Time-Series Data, Multivariate Behav. Res. 53 (2018) 453–480.
[23] D. Borsboom, M.K. Deserno, M. Rhemtulla, S. Epskamp, E.I. Fried, R.J. McNally, D.J. Robinaugh, M. Perugini, J. Dalege, G. Costantini, A.-M. Isvoranu, A.C. Wysocki, C.D. van Borkulo, R. van Bork, L.J. Waldorp, Network analysis of multivariate data in psychological science, Nature Reviews Methods Primers. 1 (2021) 1–18.
[24] M. Saqr, E. Beck, S. López-Pernas, Psychological Networks: A Modern Approach to Analysis of Learning and Complex Learning Processes, in: M. Saqr, S. López-Pernas (Eds.), Learning Analytics Methods and Tutorials: A Practical Guide Using R, Springer, 2024: p. in–press.
[25] S. Epskamp, E.I. Fried, bootnet: Bootstrap methods for various network estimation routines, CRAN. R-Project. Org. (2015).