

Guest Editorial: Special Section on Learning Analytics

Dragan Gašević, George Siemens, and Carolyn Penstein Rosé

1 INTRODUCTION

THE availability of digital data affords unprecedented possibilities for analysis of different aspects of human life. These possibilities have mobilized researchers, practitioners, institutional leaders, policy makers, and technology vendors to seek ways these data can be used to understand and enhance learning and teaching. This intense interest has given rise to the formation of the new field of learning analytics. According to the Society for Learning Analytics Research (SoLAR), learning analytics is defined as “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.” The field of learning analytics was established by drawing from a wide range of areas such as learning sciences, (educational) data mining, language technologies, machine learning, information visualization, psychology, and educational theory.

Early research publications in the field of learning analytics offer much promise in different aspects related to learning, teaching, and education. Notable examples include identification of learners at risk of failing or dropping out of a course, understanding of information flows in social interactions, or identification of different cognitive, metacognitive, and affective states in discourse and other traces of interaction with (learning) technologies. These results have attracted many institutions to invest in systemic implementation of learning analytics, development of relevant institutional policies, and creation of partnerships with organizations specializing in learning analytics. In spite of this promise, the field is in its energetic youth, and there are numerous questions that provide opportunities for growth as the field seeks to unlock the full potential of learning analytics.

This special section was organized with an understanding of the state-of-the-art in learning analytics related to

- methods that can identify the factors that influence learning optimize learning outcomes and processes across cognitive, metacognitive, social, and affective dimensions, and

- D. Gašević is with the Moray House School of Education and School of Informatics, 10 Crichton St., Edinburgh EH8 9AB, United Kingdom. E-mail: dragan.gasevic@ed.ac.uk.
- G. Siemens is with the LINK Research Lab, University of Texas Arlington-Box 19012, Arlington, TX 76019-0012. E-mail: gsiemens@uta.edu.
- C. P. Rosé is with the Language Technologies Institute and HCI Institute, Carnegie Mellon University Gates-Hillman Center 5415, 5000 Forbes Ave., Pittsburgh, PA 15213-3891. E-mail: cprose@cs.cmu.edu.

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below. Digital Object Identifier no. 10.1109/TLT.2017.2670999

- theoretically sound and empirically validated frameworks for presentation of results to users, personalization, systemic adoption, ethical use of data, and privacy protection.

2 PAPERS IN THE SPECIAL SECTION

The response to the call for papers for the special section was over 40 submissions, out of which six made the final cut for publication. Generally, the papers can be grouped into three main themes—prediction of academic performance, learning analytics visualizations, and privacy protection.

The first theme is focused on the prediction of academic performance and consists of two papers. In the first paper, “Towards Actionable Learning Analytics Using Dispositions,” Dirk T. Tempelaar, Bart Rienties, and Quan Nguyen propose the use of dispositional data (e.g., affective states and motivation) as a way to advance understanding of the association between digital traces of user interaction and academic performance. Tempelaar and his colleagues use several well-established self-report instruments for measurement of relevant dispositions along with trace data and data from student information systems. Their study shows significant advances in the predictive power of models when disposition data are used. In the second paper “Predicting Student Performance from LMS Data: A Comparison of 17 Blended Courses Using Moodle LMS,” Rianne Conijn, Chris Snijders, Ad Kleingeld, and Uwe Matzat report on the findings of a study that replicated the methodology and theoretical model proposed in [1]. Specifically, the study used data from 17 courses to confirm the importance of accounting for instructional conditions when predicting academic performance from trace data. Both Tempelaar et al. and Conijn et al. studies emphasize the importance of identification of theoretically and pedagogically-relevant predictors, which are both based on dispositions and specific to a given course, respectively, to inform generation of actionable feedback.

The second theme includes three papers and is dedicated to issues related to the visual presentation of learning analytics results to users. The first paper in this group, “Perceiving Learning at a Glance: A Systematic Literature Review of Learning Dashboard Research,” authored by Beat A. Schwendimann, María Jesus Rodríguez-Triana, Andrii Vozniuk, Luis P. Prieto, Mina Shirvani Boroujeni, Adrian Holzer, Denis Gillet, and Pierre Dillenbourg, is a systematic review of learning analytics dashboards published between 2011 and 2015. The authors organize their results around learning contexts that informed the development of dashboards, the types of

dashboards developed, and evaluation methods used. In the second paper “Widget, Widget on the Wall, Am I Performing Well at All?,” Maren Scheffel, Hendrik Drachler, Joop de Kraker, Karel Kreijns, Aad Sloomaker, and Marcus Specht report on the findings of a study in which they developed a dashboard in the form of a widget that promotes group awareness. Their empirical findings suggest that grades and widget indicator scores are positively correlated. Scheffel et al. posit that this correlation can be a useful foundation for the development of guidelines for visualization interpretations by students and tutors. In the final paper in this group “A Novel Web-Based Approach for Visualization and Inspection of Reading Difficulties on University Students,” Carolina Mejia, Beatriz Florian, Ravi Vatrpu, Susan Bull, Sergio Gomez, and Ramon Fabregat present an approach to the visualization and inspection of reading difficulties by building on the principles of open learner models. The authors also explore different opportunities for personalization of the visualization by making use of personal details, reading profiles, preferences, and cognitive traits. The empirical findings of the Mejia et al. study show a potential of the proposed visualization approach to promote student reflection and self-regulation.

The final theme examines issues of privacy protection and includes the paper “Privacy-Preserving Learning Analytics: Challenges and Techniques” authored by Mehmet Emre Gursoy, Ali Inan, Mehmet Ercan Nergiz, and Yucel Saygin, arguing that learning analytics research should make use of advanced privacy protection methods that offer quantifiable and formal mathematical proofs. With this in mind, the authors present a proof-of-concept implementation of privacy protection tools and an experimental validation of the tools on synthetic datasets.

3 DISCUSSION

To provide a systematic assessment of the current state of the field, based on the selected papers, we used the consolidated model of the field of research and practice that is proposed [2]. We analyze the papers across the three mutually connected dimensions of learning analytics—theory, design, and data science—that are suggested to be considered in learning analytics research and practice.

3.1 Theory

The importance of theoretical grounding (in contrast to atheoretical data-driven approaches) is now commonly recognized in learning analytics research and practice [3]. Theory is important to guide choices of data [4] as pertinent proxies of learning processes [5]. Theory can also guide the inclusion of relevant contextual factors and explain replicability of the results or the lack thereof (e.g., differences in factors predictive of learning gains) [1]. The papers included in the special section made use of a wide range of theoretical frameworks. Tempelaar and his colleagues use expectancy-value theory, motivation, and the engagement wheel framework, achievement goal orientation theory, self-regulated learning, control-value theory of achievement emotions, and epistemic emotions. Conijn et al. and Scheffel et al. build on theories of self-regulated learning to explain the factors that predict learning performance and guide

learners’ decisions. Finally, Mejia et al. build on psychological research on reading difficulties and activity theory to inform dashboard design.

There are several opportunities that we suggest where further integration of learning analytics and theory is promising. First, rather than only making use of existing theory to inform the development/application of learning analytics, learning analytics studies should also explore the ways to inform existing theories while also building new ones. This is expected if learning analytics aims to advance research and practice of learning, teaching, and education [2], [3].

Second, the design of learning analytics dashboards should be guided by learning theory. As the Schwendimann et al. systematic survey revealed, there is only a small fraction of studies that significantly engage a learning theory to inform the design of learning analytics dashboards. The use of principles from interaction design and human-computer interaction is certainly needed but insufficient for the field that seeks to understand and optimize learning. The approach based on activity theory followed by Mejia et al. is a promising direction, though more explicit use of learning theory would offer even more promise.

Finally, the development of sophisticated algorithms as done by Saygin et al. for privacy protection of learning analytics can benefit from relevant theory. The use of relevant theories from the areas such as sociology, philosophy, and law can offer a sound foundation to critically interrogate relevance, implications, and validity of the proposed computational methods.

3.2 Design

Design is the core of learning analytics research and practice and relates to interaction and visualization design, learning design, and study design [2]. Most of the studies presented in this special section included at least some of these design dimensions, but mostly those related to study design and interaction/visualization design. Learning design is only explicitly acknowledged in the Conijn et al. study on the prediction of academic performance based on trace data. Some of the authors of other papers included in the special section have considerations of learning design in their other studies [6].

Consideration of learning design is critical for an effective development and use of learning analytics. While a good visualization can offer much insight into different patterns emerging from data, if the visualization is not designed to support a particular (set of) learning tasks, the effects of visualizations can be negligible and even negative in some cases [3]. Attention to learning design is also essential to understand the extent to which some analytical/predictive methods and visualizations can be reused as investigated in the Conijn et al. study.

3.3 Data Science

The collection of unprecedented amounts of data about learning and learning contexts is one of the key premises on which the field of learning analytics has emerged. The use of methods from machine learning is frequently assumed in reference to learning analytics. The special

section, however, shows that a majority of the selected papers used methods for (inferential) statistics such as correlation analysis, regression models, and structural equation modeling. Only Saygin et al. go into methods that are beyond conventional statistics, while Tempelaar et al. report on the use of a clustering algorithm.

Although the use of the above methods is certainly welcome in learning analytics, adoption and advancement of the state of the art methods from machine learning and data mining should also be stressed. The study of methods that can account for the temporal and non-linear nature of learning is of particular significance [7]. Such methods can enable the use of data sources that can more effectively capture pertinent learning processes. The use of text mining is certainly an established direction due to the richness of discourse for revealing relevant cognitive, metacognitive, affective, motivational, and social processes [8], [9]. Further development of the methods that can analyze stream data sources such as psychophysiological measures can be an important direction if the field of learning analytics is to advance existing understanding of learning. Although it is commendable that the authors of some of the papers in this special section use self-report instruments, due to their prevalent use in foundational disciplines such as psychology, the static nature of such instruments is insufficient to account for the dynamic nature of learning and relevant learning process [10].

4 SUMMARY

We hope the papers in the special section will offer an informative snapshot of the current state in learning analytics. We anticipate that these papers will spark further dialogue among those interested in learning analytics. Finally, we hope that the papers included in the special section and this editorial surfaced some of the complexity of the field of learning analytics and critical roles of researchers and practitioners from a multitude of disciplines play. Learning analytics is not just a field of those from a single discipline such as learning science, education, psychology, human computer interaction, philosophy, machine learning, sociology and many others. Learning analytics as a bricolage field can only be successful if all these diverse and complementary voices are heard, contrasted, and working together.

ACKNOWLEDGMENTS

We wish to thank to numerous colleagues who dedicated their time to review the papers submitted to the special section. We would also like to thank the editorial team of the *IEEE Transactions of Learning Technologies* for their tireless support throughout the entire process.

REFERENCES

- [1] 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 learning success," *Internet. High. Educ.*, vol. 28, pp. 68–84, 2016.
- [2] D. Gašević, V. Kovanović, and S. Joksimović, "Piecing the learning analytics puzzle: A consolidated model of a field of research and practice," *Learn. Res. Pract.*, vol. 3, no. 2, 2017.
- [3] 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.
- [4] A. Wise and D. W. Shaffer, "Why theory matters more than ever in the age of big data," *J. Learn. Anal.*, vol. 2, no. 2, pp. 5–13, Dec. 2015.
- [5] P. Reimann, "Connecting learning analytics with learning research: The role of design-based research," *Learn. Res. Pract.*, vol. 2, no. 2, pp. 130–142, Jul. 2016.
- [6] B. Rienties and L. Toetel, "The impact of learning design on student behaviour, satisfaction and performance: A cross-institutional comparison across 151 modules," *Comput. Hum. Behav.*, vol. 60, pp. 333–341, Jul. 2016.
- [7] I. Molenaar, "Advances in temporal analysis in learning and instruction," *Frontline Learn. Res.*, vol. 2, no. 4, pp. 15–24, 2014.
- [8] R. Azevedo, "Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues," *Educ. Psychol.*, vol. 50, no. 1, pp. 84–94, 2015.
- [9] I. K. Howley and C. P. Rose, "Towards careful practices for automated linguistic analysis of group learning," *J. Learn. Anal.*, vol. 3, no. 3, pp. 239–262, Dec. 2016.
- [10] M. Zhou and P. H. Winne, "Modeling academic achievement by self-reported versus traced goal orientation," *Learn. Instr.*, vol. 22, no. 6, pp. 413–419, Dec. 2012.



Dragan Gašević is a professor and the chair in learning analytics and informatics in the Moray House School of Education and the School of Informatics at the University of Edinburgh. As the president (2015-2017) and a co-founder of the Society for Learning Analytics Research, he has had the pleasure to serve as a founding program co-chair of the International Conference on Learning Analytics & Knowledge (LAK) in 2011 and 2012, general chair of LAK in 2016, founding program co-chair of the Learning Analytics Summer Institute (LASI) in 2013 and 2014, and a founding editor of the *Journal of Learning Analytics*. A computer scientist by formal education, Dragan considers himself a learning analyst whose research centers on self-regulated and social learning, higher education policy, and data mining. He is a frequent keynote speaker and a (co-)author of numerous research papers and books.



George Siemens is a professor and executive director of the LINK Research Lab at the University of Texas, Arlington, and holds a faculty cross-appointment with Athabasca University. He is the founding president of the Society for Learning Analytics Research and served as the chair of the first LAK conference in 2011. He is an internationally renowned author, researcher, and theorist in the field of learning, knowledge management, and technology. He has organized and presented numerous open online courses

including two entitled *Connectivism* and *Connective Knowledge*. He is a highly sought consultant and keynote presenter, having shared his expertise in over 30 countries.



Carolyn Penstein Rosé is a professor of language technologies and human-computer interaction in the School of Computer Science at Carnegie Mellon University. Her research program is focused on better understanding the social and pragmatic nature of conversation, and using this understanding to build computational systems that can improve the efficacy of conversation between people, and between people and computers. She serves as past president of the International Society of the Learning Sciences,

executive editor of the *International Journal of Computer Supported Collaborative Learning*, and associate editor of the *IEEE Transactions on Learning Technologies*.