Supporting higher education students through analytics systems
Guest Editor: Dirk Ifenthaler

1 Guest editorial
4 Dual perspectives on learning analytics in higher education
Andrea Harkins Parrish and Laila J. Richman
15 Four design principles for learner dashboards that support student agency and empowerment
Liz Bennett and Sue Folley
27 Leveraging learning analytics for student reflection and course evaluation
Devrim Ozdemir, Heather M. Opseth and Holland Taylor
38 Engaging online students through peer-comparison progress dashboards
Paula Smith
57 Framework to support personalized learning in complex systems
Aklilu Tlahun Tadesse and Pål Ingebrigt Davidsen
86 What do first-year students need? Digital badges for academic support to enhance student retention
Dana-Kristin Mah and Dirk Ifenthaler
97 Enrollment patterns and students’ risk of academic difficulty
Robert Matthew DeMonbrun, Michael Brown and Stephanie D. Teasley
109 Making good suggestions in analytics-based early alert systems: shaping minds and changing behaviours
Dennis Foung
124 Drop-out in programming courses – prediction and prevention
Anja Haußlitschek, Veit Köppen, André Dietrich and Sebastian Zug

ISBN 978-1-83982-528-6

www.emeraldinsight.com/loi/jarhe
Supporting higher education students through analytics systems

Learning analytics and (big) data in higher education are emerging topics, however, empirical evidence as well as organisation-wide implementation are still scarce (Ifenthaler, 2017). Research on student retention has been conducted predominantly in English-speaking countries such as Australia, UK or USA (e.g. Bean, 1982; Krause et al., 2005; Mah et al., 2019; Tinto, 1993). Findings highlight that students may benefit from analytics systems through personalised and adaptive support during their learning journey (Ifenthaler, Yau and Mah, 2019). For example, students often enter higher education academically unprepared and with unrealistic perceptions and expectations of academic competencies for their studies. Both, the inability to cope with academic requirements as well as unrealistic perceptions and expectations of university life, in particular with regard to academic competencies, are important factors for leaving the institution prior to degree completion (Mah and Ifenthaler, 2018).

Analytics systems for supporting learning and teaching in higher education are slowly moving towards a mature field of research and development (Ifenthaler, Mah and Yau, 2019; Schumacher and Ifenthaler, 2018a, b; Schumacher et al., 2019). This broader (and system wide) adoption of analytics system provides new testbeds for empirical research and areas of new discoveries for the learning as well as data sciences (Mah et al., 2019).

This special issue brings together the scholarly research and theory focusing on contemporary issues related to analytics system and how they support students’ at higher education institutions. The contributions provide insights into how educational data, analytics systems and advanced digital technologies contribute towards successful learning and teaching scenarios at higher education institutions.

Paper selection process

In summer 2018, a call for submissions was circulated through electronic mailing lists of the following organisations: AECT (Association for Educational Communications and Technology), AERA (American Educational Research Association), ASCILITE (Australasian Society for Computers in Learning in Tertiary Education), as well as through the regular channels of educational technology groups. The call defined the focus of the potential submissions as follows:

- Do analytics data change student behaviour (e.g. learning strategies) and dispositions (e.g. motivation, emotion)?
- Understanding the learning journey of higher education students.
- Leveraging assessments analytics for personalised feedback.
- Linking analytics data and study success.
- Building evidence for graduate analytics.
- Change management for data analytics at higher education institutions (e.g. implementation strategies, requirements).

The guest editor of this special issue is very thankful for all support received from Editor of Journal of Applied Research in Higher Education, and the reviewers who ensured the quality of this volume.
Initially, 16 abstracts were submitted by the end of September 2018. Upon careful review and agreement among the advisory board, nine of them were invited to submit a full manuscript by the end of November 2018. Main criteria for the selection of manuscripts was a clearly articulated focus on analytics systems for supporting students in higher education and how well this focus was consistently enunciated throughout the proposed work. Each manuscript was assigned to at least three reviewers of the special issue review board and additional reviewers for the *Journal of Applied Research in Higher Education*. All of the initial reviews were completed by the end of March 2019. Based on the comments of the reviewers and on the individual feedback of the guest editor, the manuscripts were moved to the second round of reviews. Authors were asked to submit their revised manuscript by the end of May 2019 addressing the comments from the reviewers and from the guest editor. The final acceptance of manuscripts was completed by the end of June 2019.

**Contributors to this special issue**

This special issue starts with a research paper by Andrea Parrish and Laila Richman, “Dual perspectives on learning analytics in higher education”, which suggests perspectives on learning analytics from an administrator and a faculty member. The dual perspectives help to further the dialogue between university administration and faculty as analytic systems become more widespread.

Liz Bennett and Sue Folley present in their research paper, “Four design principles for learner dashboards that support student agency and empowerment”, a student-centred perspective to understanding the range of ways that students respond to receiving information about their learning behaviours presented on a dashboard. The four principles are: designs that are customisable by students; foregrounds students’ sense making; enables students to identify actionable insights; and dashboards are embedded into educational processes.

“Motivating online students through peer-comparison progress dashboards” is a practitioner report by Paula Smith sharing student and tutor perspectives on the use of dashboards to increase online students’ motivation, and it examines whether the benefits of a peer-comparison dashboard are reserved for high-achieving students.

The practitioner paper, “Learning analytics for student reflection and course evaluation”, by Devrim Ozdemir, Heather Opseth and Holland Taylor demonstrates a process of faculty utilisation of learning analytics by evaluating students’ course objective achievement results to enable student reflection, student remediation and faculty curriculum evaluation. Learning analytics enabled meaningful conversations focusing on course learning objectives and provided detailed information on each student. The learning analytics tool also provided detailed information regarding which areas faculty needed to improve in the curriculum.

Aklilu Tilahun Tadesse and Pål Davidsen present in their research paper, “Framework to support personalized learning in complex systems”, a design framework applied to the creation of a personalised and adaptive online interactive learning environment. Findings suggest that the use of personalised and adaptive learning tools support learning about complex systems.

“What do first-year students need? Digital badges for academic support to enhance student retention” is a research paper by Dana-Kristin Mah and Dirk Ifenthaler investigates data on first-year students’ needs regarding academic support services and reasons for their intention to leave the institution prior to degree completion. It is suggested that higher education institutions can create digital badge programmes, which may improve communication of academic requirements and may also serve as a platform for a staff-student conversation about expectations and demands for a successful first-year experience.

The research paper by Robert DeMonbrun, Michael Brown and Stephanie Teasley, “Enrollment patterns and students’ risk of academic difficulty”, expands upon an existing framework that investigates students’ academic difficulty in co-enrolled courses by adding additional co-enrolment variables that may influence academic performance in introductory
gateway courses. The authors provide a generalisable methodology that can be used by other institutions to investigate curricular pathways that have the potential to increase study success.

“Shaping minds without changing behaviours: predictive analytics in a university english course” is a research paper by Dennis Foung which investigates students’ perceptions of the course diagnostic reports tool and whether they plan to act on the recommendation. The findings provide useful insights for early alert system designers to establish a system for generating practical recommendations for students.

The final research paper, “Dropout in programming courses – prediction and prevention”, by Anja Hawlitschek, Veit Köppen, André Dietrich and Sebastian Zug identifies activity patterns that indicate students at risk and investigates reasons behind specific activity pattern. The findings indicate a link between activity patterns and learner characteristics. Instructional interventions to support students and to prevent dropouts are suggested.

The nine papers cover a wide range of contributions focussing on supporting higher education students through analytics systems and provide empirical evidence as well as practical implications for an emerging field in higher education research. The theoretical foundations, insightful findings and innovative frameworks, as well as practitioner reports shall inspire future high-quality research studies and contribute to the growing knowledge base of educational technology, learning analytics, and higher education learning and teaching practice.

Dirk Ifenthaler

Department of Learning, Design and Technology, Fakultat fur Betriebswirtschaftslehre, Universitat Mannheim, Mannheim, Germany

References
Dual perspectives on learning analytics in higher education

Andrea Harkins Parrish
Department of Special Education,
Towson University, Towson, Maryland, USA, and
Laila J. Richman
College of Education, Towson University, Towson, Maryland, USA

Abstract
Purpose – In higher education, the authors serve multiple stakeholders with varying perspectives on the institution’s learning analytic system. The purpose of this paper is to highlight dual perspectives within the learning analytics (LA) system in one institution: that of an administrator responsible for college-wide improvement, and a faculty member responsible for programmatic improvement.

Design/methodology/approach – This manuscript provides a critical perspective with dual sets of experiences and viewpoints. This approach allows close examination of each perspective within the context of the existing literature, as the authors examine transitions in the use of LA.

Findings – These viewpoints offer insight into the interpretation of LA through the lens of various roles. In examining these viewpoints, the authors offer three actionable steps for other institutions who seek to implement.

Practical implications – These actionable steps offer a starting point for other institutions to engage in conversations related to the adoption of LA for continuous improvement across levels and roles. Relevant implications for various parties are discussed, with an emphasis on how administrators within the university system may support faculty to incorporate LA as part of their scholarly and teaching responsibilities.

Originality/value – Few studies have examined the perspectives of multiple stakeholders within an institution. Here the authors have presented these dual perspectives in order to further the dialogue between university administration and faculty as analytic systems become more widespread. Through this dialogue, the authors see increased opportunities for all stakeholders to better understand their role in LA.

Keywords Higher education, Learning analytics

Paper type Viewpoint

Introduction
More data are being collected on college students than ever before. In fact, the sheer quantity of data being collected positions us for an “analytics revolution” (Gagliardi et al., 2018, p. 1). Now the question is: how do we learn from this data that we have collected? Siemens (2013) defines learning analytics (LA) as “the measurement, collection, analysis, and reporting of data about learners and their contexts, for the purposes of understanding and optimizing learning and the environments in which it occurs” (p. 1382). LA has only recently emerged as a topic of interest in higher education but is gaining significance as an approach that can be used to address the challenges many institutions face. Siemens (2013) encourages us to see LA as the intelligent use of data that helps us make new connections about our students so that we can better predict their success and advise them in ways that supports their achievement in the university setting.

Literature review
Learning analytics in higher education
The topic of LA is often associated with related terms, such as big data, and there is a growing body of descriptive research which compares and contrasts the various definitions of the construct. In particular, the meaning of the term big data is typically disputed by researchers (Daniel, 2017). While some focus on the more literal interpretation of the term to mean vast amounts of data collected, other see big data as our conversion to a system where
technologies are collecting more information about students’ practices and patterns over time. Daniel and Butson (2013) pose a useful conceptual framework which outlines the components of big data in higher education, and in particular, how LA fits within that framework. Daniel and Butson (2013) propose that when we use the term big data in higher education, we are really referring to four distinct systems: institutional analytics (data typically found in an institution’s data warehouse and conveyed via data dashboards); informational technology analytics (information held in the student information system and/or learning management system (LMS)); academic analytics (data collected on individual student and program performance); and LA (data collected on the contexts of students’ behaviors and activities). In this paper, we will present a critical reflection of our experiences within an institution that is transitioning to a system focused on academic analytics, while working to establish more robust systems in the other areas identified of this four-tiered framework. We also discuss the conversion to the use of academic analytics, utilizing a new digital platform.

The changing landscape of higher education
The purpose of analyzing data is for institutional leadership and faculty to make data-driven decisions to solve challenges. These challenges may range from student retention and enrollment to improved student outcomes. Institutional leaders are tasked with optimizing available resources to support these efforts, while maintaining the mission and vision of the organization. Student data support these efforts because it allows us to gain information about performance while also learning about students’ experiences, habits and patterns. When analyzed, this information allows institutions to improve services to students, namely, the ability to more accurately predict, forecast and mitigate issues related to retention, performance and improved student experiences (Bronniman et al., 2018).

LA has been used in the past to address issues related to student retention. For instance, Bronniman et al. (2018) conducted a case study in which LA was applied to the scholarship of teaching and learning. The Australian-funded project utilized LA to re-design a program that had previously experienced attrition of students from underrepresented groups. While Bronniman et al. (2018) reported that faculty were skeptical at first, it became clear as the project progressed that there were advantages to making data-based decisions regarding program changes. Bronniman et al.’s study is reflective of the growing body of work in this area.

Even with the growing body of literature, work in the area of LA is rather narrow in some areas. Viberg et al. (2018) conducted a comprehensive review of more than 250 papers on LA in higher education and these findings are useful for characterizing the benefits of LA in higher education contexts. Overall, there were a small number of studies which showed LA as a viable tool for improving student outcomes, particularly in the areas of students’ academic performance and knowledge acquisition; students’ skill development (such as time management skills); and students’ abilities to problem solve (Viberg et al., 2018). However, most of the publications on LA support its use as a tool to support teaching and learning in higher education. Viberg et al. (2018) suggested that the power of LA lies in how we utilize various forms of data to support student retention, such as methods to identify student readiness; to identify areas in which students need support or even to target learner preferences or habits that affect their ability to progress satisfactorily in their studies. These findings suggest that LA can; therefore, be a tool that institutions and its faculty apply to support the learning outcomes of students.

Technology’s role in learning analytics
Digital tools are becoming more sophisticated and are now able to capture more information about an institution and its learners (Daniel, 2017); and this rise in LA is driven, in part, by our
new technological capacities to collect and analyze large sets of data (Bronniman et al., 2018). For example, data dashboards and academic analytic systems provide greater capabilities to collect and analyze data across institutional systems. These trends, discussed in the literature, are supported by the 2018 higher education edition of *The Horizon Report*, which names analytics technologies as an accelerating trend in higher education (Adams Becker et al., 2018). The development of new technologies, such as assessment management systems (AMS) designed to provide academic analytics, draw us away from static, paper-based metrics and toward systems which provide real-time data to inform faculty decisions. However, conversion to a digital analytics system requires changes in the way data are collected and analyzed at all levels within an institution (Adams Becker et al., 2018). In the example presented within this paper, we will address some of our own experiences related to the conversion from a traditional (paper-based) analysis of data to a digital collection and analysis of data facilitated by a new AMS.

*Academic faculty perceptions of learning analytics*

W. Edwards Deming (1942) said:

> [...] data are not taken for museum purposes; they are taken as a basis for doing something. If nothing is to be done with the data, then there is no use in collecting any. The ultimate purpose of taking data is to provide a basis for action or a recommendation for action. (p. 173)

If rich data exist, will our academic faculty choose to use it or reject it? Svinicki et al. (2016) sought to answer that question by looking at the relevant factors associated with faculty members’ participation in the LA process. They found that it was the faculty members’ self-efficacy in data collection and analysis that was the most important factor in their participation. Faculty also needed to believe in the utility of the data in order to expend effort in using it to make instructional improvements (Svinicki et al., 2016). Not surprisingly, Svinicki et al. (2016) found that whether a faculty member utilized available data was largely dependent on his or her beliefs as to its benefits. If an individual’s positive beliefs regarding the potential benefits outweighed the perceived effort of time and energy, the faculty member was more likely to become engaged in the LA system that was established in the institution.

Additional studies on faculty members’ perceptions of LA have been conducted. Howell et al. (2017) examined academic faculty members’ perceptions on LA and identified several main themes following focus groups with 35 higher education faculty members. Findings indicated that faculty saw the potential for learner analytics to benefit teaching and learning as well as to identify early on when students needed support (Howell et al., 2017). However, there were concerns among faculty as to their obligations for data analysis within an already overburdened workload. Faculty worried about their ability to provide all of the interventions, without additional resources. Faculty members also raised concerns that the data would not be used to improve practice or that there was potential for misuse, including inappropriate usage, the misinterpretation of data or the risk to students’ privacy (Howell et al., 2017). Overall, results showed that academics’ perspectives toward LA had somewhat of a push-pull syndrome. Faculty members saw the benefits of LA from an “economic imperative,” such as the individual achieving success as a college student but they also had concerns about the “moral imperative” surrounding the use of data and how it would be used (Howell et al., 2017, p. 9). Some faculty struggled with the idea that the institution had collected large quantities of data that they needed to utilize, rather than a targeted collection of information for a specific purpose. Herein lies the crux of the issue from the faculty perspective: depending on how it is presented, LA may appear to be an add-on. However, when LA is instituted as an integrated approach that very specifically targets improvements to both teaching and learning, faculty see the value in the approach.

Research indicates that faculty involved in LA initiatives wish to use the information they gain to improve their instructional practice. In Kahn’s (2017) qualitative examination of
academic faculty members’ perceptions, participants indicated that any data that could be used to tailor their teaching practices or increase the engagement of their students would be useful (Kahn, 2017). While these faculty did not see the need for all forms of data related to student activity, they were interested in having access to the information that would help them to make necessary pedagogical changes. Similarly, Knight et al. (2016) found that academic faculty commonly used student performance data to adjust the courses they were teaching or to improve their instructional methods. Faculty shared their positivity regarding the usefulness of information that allowed them to re-calibrate syllabi, assignments or the pedagogy used to support students in learning course content (Knight et al., 2016).

Challenges and ethical issues
LA frameworks should be implemented with an explicit focus and for an intended purpose (West et al., 2016). Concerns surrounding the ethical use of student data is not only a concern of faculty (as indicated in Howell et al., 2017), but can also be of concern to students (Roberts et al., 2016).

When discussing the use of student data with academic faculty, Knight et al. (2016) found that faculty advocated a cautionary approach to using data in a way that could potentially stereotype students. For instance, while the intent could be to identify students in need of academic support, sharing this with a new instructor or advisor could label that individual as a “bad student.” In this way, advanced warning systems designed to identify and support, may have unintended consequences that result in negative outcomes or unnecessary bias toward students. In a large-scale survey of administration, faculty and support staff views on LA across 22 institutes of higher education, West et al. (2016) found stakeholders’ chief concern to be student privacy, confidentiality and their right to informed consent. Faculty members likened these ethical issues to the standard of ethics required in conducting research and felt that collecting any data without a pre-specified purpose would violate students’ rights (West et al., 2016).

Inquiry into students’ perceptions of LA also echoes these same ethical concerns. Concerned by a lack of student voice in the discourse, Roberts et al. (2016) conducted a qualitative inquiry into students’ perceptions of LA. First, Roberts et al. (2016) found that students were initially unaware of the term LA or the amounts of data that were being collected about them by their institution. As they learned more, students discussed both advantages and concerns. They saw advantages to the use of data for academic support; viewed LA as a potentially useful way to stay motivated; and some felt positively about the opportunity to receive ongoing notifications regarding their progress (Roberts et al., 2016). However, students had concerns about their privacy and the potential for bias among faculty. They questioned how the institution could be certain that all the data were accurate and expressed general concern about being compared to their peers. Students also provided examples of how repeated notifications regarding their performance could potentially impede their independence (Roberts et al., 2016).

The implementation of a cultural shift in a sector as complex as higher education is likely to come with challenges. Macfadyen et al. (2014) suggest that one way to foster cohesiveness in the midst of such change is the installation of institutional policy that explains and supports the role of LA. Given the emergence of research in this area, there is now a growing body of literature available to help institutions and statewide systems develop LA policies which are rooted in empirical research.

A critical reflection with dual perspectives
While many institutions collect and analyze data related to student performance, practitioners’ voices have largely been left out of the LA discourse (Feng et al., 2016). Therefore, in this critical reflection we will highlight multiple stakeholder experiences and
perspectives related to LA within the context of the literature. In doing so, we will discuss LA at both the micro and macro levels within one institution. This approach is in alignment with Siemens' (2013) early recommendations on the rise of LA in higher education. Siemens (2013) suggested that LA in education can exist at various levels, ranging from the micro (individual classroom, cohort, program or faculty member) to the macro level (university, state system or region). In this reflection, we will examine these contexts by focusing on the perspectives of an administrator and a faculty member within a large, mid-Atlantic university in the Northeastern USA. This university enrolls approximately 23,000 students per year, primarily at the undergraduate level and celebrates a diverse campus with almost 50 percent of its student body being diverse. The university is organized into six discipline specific colleges and an Honors College, employing approximately 1,700 faculty. Within this faculty count, it is important to note fewer than half are tenured/tenure-track faculty.

The College of Education (COE), more specifically, oversees all teacher preparation and certification programs at the University. The College includes six departments: Elementary Education, Early Childhood Education, Special Education, Instructional Leadership and Professional Development, Educational Technology and Literacy and Secondary Education. Within each department are several graduate and/or undergraduate programs that align with teacher certification areas. In 2019, the COE enrolled approximately 2,300 students and employed 128 full-time faculty and 94 part-time faculty.

The administrative structure of the COE includes a dean, associate dean, assistant dean, department chairpersons and program coordinators. Program coordinators in collaboration with the department chairpersons are primarily responsible for program level assessment and these individuals write annual reports that document their program data, the analysis of that data and a strategic plan for program improvement. The associate dean is responsible for assessment across the College and reviews all program reports which are used to support the College's strategic plan. While both the administrator and the individual faculty members have a background in instruction and assessment as part of previous teacher preparation programs, the COE has not instituted specific training or faculty development in LA to date. Faculty have been trained in the use of the AMS to enter student data and are in various stages of implementing additional features of the AMS throughout the College.

The administrative perspective
The COE reports to several accrediting bodies in order to offer teacher certification programs. At the highest level, the COE participates in institution level assessment targeting the requirements of our regional accreditor. Discipline focused accountability includes national accreditation by the Council for the Accreditation of Educator Preparation, content specific recognition by Specialty Program Associations (e.g. National Council of Mathematics, Council for Exceptional Children, etc.), and program approval by the state. If the review criteria of any one of these entities are not satisfied, the COE could lose the ability to prepare and graduate education professionals. While the requirements of these oversight bodies typically overlap, there are often still unique aspects and standards that need to be addressed. The delicate balance is addressing these external accountability requirements while also engaging in LA focused on improving learner outcomes.

In 2013, the National Institute for Learning Outcomes Assessment conducted a survey of over 1,200 provosts nationwide. One of the major findings of the survey was that compliance with accrediting agencies was identified as the primary impetus for assessment activities (Kuh et al., 2014). Kuh et al. (2015) call on higher education to use assessment as evidence for improving student outcomes instead of perpetuating a “culture of compliance.” Jankowski et al. (2018) found in the most recent survey of provosts that while the focus on assessment for program improvement has increased, accreditation is still the leading driver...
of analysis of student performance. Having a high-quality and authentic assessment system requires that assessment be a part of instruction and not an “add-on” or “separate” process completed merely for compliance reasons. Kuh et al. (2015) posit that assessment should be “actionable, focused on the needs and interests of end users, embedded in the work of teaching and learning, available in understandable forms, customized, and supported by institutional leaders” (p. 9). The use of LA supports this authenticity (Kahn, 2017) in order to promote faculty engagement in the process.

In line with this concept of assessment, it is important that any assessments used for accreditation purposes be embedded within program coursework and driven by faculty. In order to accomplish this goal, faculty work collaboratively at the program level to identify existing or develop new assessments that provide evidence of mastery of content program standards (Howell et al., 2017). These assessments are then administered through the AMS, allowing faculty to assess student work directly within the system where the data is captured and analyzed. Within this AMS, scoring rubrics are “tagged” with the relevant program level outcomes as well as any relevant state and national indicators. This tagging allows programs to analyze the vertical alignment of learning outcomes across these various levels. The AMS also provides faculty with real-time analytics on how students within their course are performing on the various rubric criteria and related standards. At the program level, faculty can look at LA across all course sections to get a broader perspective on student outcomes from each assessment.

Moving beyond the individual program level, the COE has also established cross-departmental faculty workgroups to identify and/or develop assessments that cut across all programs to measure broader pedagogical outcomes. In order to do this, while still allowing autonomy at the program level, “core” scoring rubrics, which measure critical teaching concepts/performances regardless of discipline, were developed. While each program is expected to use the same “core” scoring rubrics, care was taken to allow each program flexibility in the creation and implementation of the learning activity itself as well as the ability to add to the scoring rubric if needed. This approach of focusing on the evaluation of the competencies/outcomes instead of a particular process or product preserves the ability to engage in assessment that is universal across programs and provides a common measure, while still an authentic part of teaching and learning. Data from these “core” scoring guides allow for analysis at the college level and, in some cases, at the national level.

The next critical step after analyzing the data is having faculty share and discuss results with a focus on identifying and implementing any necessary changes (Kinzie et al., 2015). While this step happens at the individual program level, it is also done at the college and institution level. For the COE, an annual report, including analysis, recommended changes and measures of the impact of the previous year’s changes are reported. These reports are then shared at an institution-wide “Assessment Day,” where faculty reflect on their own program assessment results while also providing feedback and suggestions to others. This provides faculty with an opportunity to engage in the work of assessment in ways that move beyond compliance and toward a “culture of evidence” (Shavelson, 2007).

While the process that has been developed relies heavily on advances in technologies such as the AMS and its ability to tag learning outcomes to rubrics and on-demand, interactive analytics, it has become apparent that transitioning away from previous processes is difficult for faculty. Some faculty struggle with giving up printed tables and spreadsheets in exchange for dynamic online reports that allow the user to “dig” deeper into the data, drilling as far down as specific student artifacts. For example, if the interactive report generated by the AMS shows four students scored low on one rubric criteria, the faculty can click on that rubric cell to see the actual work submitted and faculty feedback provided for each of those four students. While faculty still value the analysis process, they do not yet understand the power of the technology and how it can allow them to engage
more acutely with the data. Some faculty struggle to assess in the AMS, opting to grade on paper and then transfer their results into the system, creating more work and an “add-on” component for themselves. Supporting faculty in transitioning from traditional “assessment” to more dynamic LA will require continued and strategic support.

The faculty member perspective
A faculty member’s responsibilities within the institution are typically related to teaching, scholarship and service. In particular, faculty members engage in data-based decision making using the adopted LMS, AMS or direct observations of students’ performance. For instance, faculty serving as program coordinator access course evaluations of faculty members and student performance data in the LMS and AMS. Program coordinators are also often required to monitor program data by conducting analyses of enrollment, recruiting, retention and program-wide performance data (such as selected standardized assessments within a program of study). As the amount and types of student data proliferates, faculty members need to have access to the types of information that can inform their teaching and improve students’ learning. They may also benefit from data which can be used to support their scholarly or service-oriented efforts at the university. Additionally, program coordinators need information that help them to work with teams of faculty to create a strategic and continuous cycle of improvement to the program as a whole; to individual courses and course requirements; and to address issues related to enrollment or recruitment, if applicable.

Overall, a common challenge faculty experience is the ongoing reality of balancing these responsibilities. As Macfadyen et al. (2014) write, “It is unrealistic to consider that educators will adopt time-consuming longitudinal and personalized assessment models given the massive increase in resources and workload that would be acquired” (p. 21). If faculty and administration agree with that sentiment, how can LA become a part of the ongoing review of student performance and program performance reviews that are likely happening in many institutions? Therefore, we have found that opportunities to engage faculty in LA that is directly related to their individual teaching, scholarship or service has been the most fruitful. Bronniman et al. (2018) write specifically on this issue and encourage institutions and individual faculty members to use LA processes as a way to focus on improving the pedagogy in higher education classrooms or to utilize the data in ways that creates a more solid evidence base.

The concept of academic freedom among university professors is a long-held virtue in the field of academia. In particular, individuals with the highest levels of technical expertise in their area of study expect the opportunity to make their own legitimate judgments about the content, their teaching, and what their students need in order to be successful in higher education coursework. As part of this process, there tend to be sources of informal data that may impact faculty members’ opinion. For instance, after reviewing the performance on a particular course requirement, a professor may choose to clarify the instructions or change his or her pedagogy related to the learning objective. But these ongoing forms of data-based decision making are often not accounted for in the LA system. Therein, we see the tensions between academic freedom and LA arise.

Overall, the transition to a cohesive and integrative approach to LA at our institutions is still occurring. We have transitioned to a digital AMS for the collection of student data and have begun working with faculty on methods for analyzing and interpreting student data in this digital platform. We have also begun to build systems for using our AMS in more dynamic ways, such as creating a wholistic “picture” of each student’s progression through their program including admissions data, advising data, internships and assessment data. We have worked to ensure faculty, advisors and students have access to the data and understand how it can be used to support teaching and learning. As the literature on
LA in institutions like ours has grown, we have learned much from other institutions who are undergoing similar conversions and in sharing our perspectives, we hope to support others working toward developing a similar LA culture that uses data to support both students and the faculty who serve them.

**Practical implications**

In order to better facilitate higher education’s entry into the “analytics revolution” (Gagliardi et al., 2018, p. 1), institutions have the opportunity to support change. Building a culture of data-based decision making through LA involves a coordinated effort to develop shared goals, build capacity for LA within the institution and to widen the focus beyond accountability to continuous improvement. In this section, we provide implications for facilitating change that are based upon our highly contextualized experiences within our own institution. Therefore, we present these as factors to be considered by other administrators and faculty members not as a recipe for adoption, but rather as considerations to guide dialogue about change. Every institution has its own complex ethos and we encourage readers to determine the potential relevance to their own context.

*Develop shared goals*

A key component of success in LA is the co-development of goals clearly outlining what faculty want to achieve as well as what the program, college and institution wish to gain (Bronniman et al., 2018). Coupled with this, institutions can establish clear and consistent terminology around LA within the institutions and then discuss the parameters for addressing the aforementioned potential ethical issues inherent in student data usage. Utilizing the framework developed by Daniel and Butson (2013) may be helpful for institutions in identifying the various types of LA within the organization. Current research highlights the need to develop scholars with interdisciplinary skills in data analytics, technology and learning to further solidify this cultural shift (Daniel, 2017; Ifenthaler, 2017). This blended and engrained knowledge and understanding of LA serve to sustain the cultural shift and enhance capacity within institutions.

Institutions may also find it useful to establish an infrastructure for discussing and documenting knowledge from LAs at the various departmental and institutional levels. Feng et al. (2016) discuss the importance of establishing shared goals between partners engaging in LA. In order to do so, it necessary to make certain that everyone feels clear about what data are being collected and how the data will be used to inform institutional decisions. In an even more practical sense, effective administrative leaders are those who help faculty understand that both the micro and macro-level forms of data-based decision making go hand-in-hand. The micro-level knowledge that the faculty member brings to the larger framework of the LA framework is not only useful, but essential in making sound decisions with the larger sets of data we collect on student performance. Shared goals can be developed by inviting faculty members to share input at each step of the LA conversion and to honor that input is a valuable source of information in instilling change.

*Focus on continuous improvement*

An issue experienced in LA at the micro level relates to relevance. In particular, it is necessary that faculty members know what is expected of them and how their role is of value to the LA process. For instance, consider asking: how can faculty use these data to support a strategic planning process? Or, how can we ensure that the data we are asking faculty to review is well-connected with the individual’s goals in the areas of teaching and scholarly research? Making these connections, both for new and seasoned
faculty members and at the administrative level can support in the development toward continuous improvement.

“Analyzing big data requires contextual relevance and value in its utilization” (Daniel, 2017, p. 21). Likewise, it takes time for both leadership and faculty in institutions to adjust to the technological innovations that support LA, but nonetheless are new systems that require training and adjustment. Utilizing LA to improve student outcomes also requires significant collaboration, not only within programs and colleges, but across the institution (Siemens, 2013). Establishing roles and responsibilities can support the use of LA and help to transition from the traditional “bottom-up” approach to a blended approach that also incorporates “top-down” support (Siemens, 2013, p. 1391). Using continuous improvement as a framework for LA engages faculty by appealing to their desire to improve their own practice (Kahn, 2017) and this satisfies both institutional and accreditation goals. When connected in this way, LA empowers institutions to aggregate and synthesize data at multiple levels in order to use that information as the “basis for doing something” (Deming, 1942) – supporting and improving learner outcomes.

Build capacity
It is necessary for an institution to recognize the capacity of administration, faculty and support staff to conduct effective LA. Within these multiple stakeholder groups, individuals may require training, time and ongoing support to utilize student data in meaningful ways. Particularly, if looking to build a culture of LA, digital tools can be beneficial in supporting this effort. However, based on our experience, the introduction of any new technology needed to be coupled with individualized and intensive professional development for various stakeholder groups. In particular, support staff and faculty needed more than analytic strategies and benefited from instruction on the ways in which the digital platform can be leveraged to support deeper levels of analysis.

Literature within the last several years supports the notion that many organizations find themselves in similar stages of LA development and that conversion toward more sophisticated procedures for LA. Given this attention that LA has received, Ifenthaler (2017) studied higher education institutions’ capacity to conduct LA by surveying 153 higher education faculty and staff members. Overall, a majority of the institutions were identified as unprepared for LA due to a lack of adequately trained staff or access to an appropriate technology platform to support LA. Ifenthaler’s findings inform us that, as a field, we have room for growth in achieving preparedness for LA. Once we reach a greater level of preparedness, it is our role as scholars to conduct empirical research which tells us more about the impact of LA on the issues faced by our organizations.

Conclusion
There seems to be a myth that other industries have managed to address the challenges of using big data and LA to systematically address the issues they face. Further, there are mounting pressures on institutes of higher education to optimize the resources available, while addressing issues in student retention, enrollment and recruitment (Bronniman et al., 2018; Brint and Clotfelter, 2016). The growing discourse of LA in higher education is leading more institutions to claim that they are “doing it,” even though many of the various stakeholders are still identifying just exactly what “it” looks like. In this paper, we have presented what we see as an honest and transparent look at the work within one institution to grow the culture of LA within our organization. Given the literature in this area, we see other institutions sharing their ever-present need to leverage the use of analytics to drive program improvement, student success and faculty development. Zilvenski et al. (2017) provide a telling portrayal of why this can be complex to navigate: “No one has figured out the magic formula, and even very
well-known examples have not been sustained” (p. 9). But, as with any form of change, a fully engaged system of administration, faculty and students can help to shape the culture of the institution, toward data-based decision making that will ultimately improve teaching and learning.

References


Shavelson, R.J. (2007), A Brief History of Student Learning Assessment: How We got where We are and a Proposal for where to go Next, Association of American College and Universities, Washington, DC.


Corresponding author
Andrea Harkins Parrish can be contacted at: aparrish@towson.edu

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Four design principles for learner dashboards that support student agency and empowerment

Liz Bennett and Sue Folley
School of Education and Professional Development, University of Huddersfield, Huddersfield, UK

Abstract

Purpose – The purpose of this paper is to take a student-centred perspective to understanding the range of ways that students respond to receiving information about their learning behaviours presented on a dashboard. It identifies four principles to inform the design of dashboards which support learner agency and empowerment, features which Prinsloo and Slade (2016) suggest are central to ethical adoption of learning analytics.

Design/methodology/approach – The study involved semi-structured interviews with 24 final-year undergraduates to explore the students’ response to receiving dashboards that showed the students’ achievement and other learning behaviours.

Findings – The paper identifies four principles that should be used when designing and adopting learner dashboards to support student agency and empowerment.

Research limitations/implications – The study was based on a small sample of undergraduate students from the final year from one academic school. The data are based on students’ self-reporting.

Practical implications – The paper suggests that these four principles are guiding tenets for the design and implementation of learner dashboards in higher education. The four principles are: designs that are customisable by students; foregrounds students sense making; enables students to identify actionable insights; and dashboards are embedded into educational processes.

Originality/value – The paper’s originality is that it illuminates student-centred principles of learner dashboard design and adoption.

Keywords Design, Dashboards, Learning analytics, Student agency

Paper type Research paper

Introduction

Learning analytics is a rapidly growing area in higher education (Howell et al., 2018). It is defined by Slade and Prinsloo (2013, p. 1512) as the “collection, analysis, use, and appropriate dissemination of student-generated, actionable data with the purpose of creating appropriate cognitive, administrative, and effective support for learners”. Dashboards display the outputs from learning analytics in a range of visual formats for instance through graphs, bar charts and other visualisations (Khan and Pardo, 2016). The dominant use of learning analytics dashboards has been to provide data for use by academics or managers to identify students at risk of dropping out, or to identify where interventions are needed, rather than being used by students (Schwendimann et al., 2017; Willis et al., 2016). However, there is a growing interest in developing and adopting student-facing learner dashboards (Ferguson et al., 2014; Newland and Trueman, 2017).

Learning analytics involves gathering data about student behaviours that the student may not be aware of, and this leads to ethical issues related to the challenge that institutions are surveilling their students without their explicit consent (Prinsloo and Slade, 2017b). In addition, the types of data gathered, and the significance that is attached to these data make assumptions about what is valued and is shaped by those who design the algorithms (Wilson et al., 2017). Indeed, dashboard development is typically owned by computer service departments (Ferguson et al., 2014) and driven by a focus on the possibilities of large data sets to reveal new information. Frequently, students are not consulted in the design of the
Thus, providing students with access to dashboards that present them with visualisations of their data is a move that redresses the hitherto institutional focus of learning analytics.

Adoption of learning analytics and dashboards is often accompanied by a technological deterministic assumption that their introduction will inevitably lead to positive results (Howell et al., 2018). Yet Khan and Pardo (2016) found no correlation between students’ engagement with their dashboard and academic performance. Similarly, Viberg et al. (2018) based on an analysis of 252 papers on learning analytics in higher education found little evidence of improvement in student learning outcomes from the use of the analytics. Even when research has identified a positive impact on students’ motivation from use of learner dashboards, the impact is not uniform with a few students having a negative impact on their motivation (Bennett, 2017). Indeed, Bodily and Verbert who reviewed 93 papers on students’ use of learner dashboards, have noted that uptake of learner dashboards by students was low.

This means that the field needs to better understand the ways that student dashboards can be designed and implemented in order to engender positive student behaviours. This paper, therefore, sets out to establish principles for design and adoption of learner dashboards, which will help to ensure that they are effective in terms of their impact on students’ behaviours.

**Design of learner dashboards as “tools in use”**

Learner dashboards provide students with visualisations about their learning behaviour with the aim to develop “student autonomy, giving students more control over their learning and helping them feel more intrinsically motivated to succeed” (Bodily and Verbert, 2017, p. 405). There are, broadly, three types of learner dashboards: predictive, modelling and descriptive. Predictive types use machine learning based on computer algorithms which draw on a range of trace data to predict a student’s likely outcome. Frequently the algorithms are used to produce an “at risk” rating for individual students (Arnold et al., 2014; Prince, 2018). Modelling dashboards provide students with a visual representation that models their learning behaviours. For instance, aspects of a students’ online behaviours such as communication, initiative, presence that have been derived from the number of posts, the number of comments in reply to others’ posts, the time spent online, etc. The third type of dashboard is descriptive and displays past learning behaviours. Figure 1 is an example of a descriptive dashboard and depicts four dashboard elements displaying a particular student’s attendance in pie chart form, the score that they are currently on track for and their attainment in a particular assignment represented as a bar chart and in a narrative form. All three types aim to provide a student with representations of their learning data, presented in graphically rich ways.

There is a growing body of literature about the design of learner dashboards and several literature review papers which scope the field. A comprehensive literature review by Bodily and Verbert (2017) analysed 93 papers in relation to their functionality, data sources, design analysis, student perceptions and measured effects. They found that the clear majority of papers focussed on technical features such as the data sources, the functional aspects (e.g. data visualisations, what the dashboard aimed to do) whereas only two papers reported on how students’ behaviours were changed through the use of learner dashboards and only a third (34 out of 93) focussed on students’ perceptions of the design (Bodily and Verbert, 2017). Indeed, Bodily and Verbert conclude that the field needs more studies that aim to understand the students’ perspective and which go beyond the technical features to examine the dashboards as “tools in use”.

Another comprehensive meta-analysis examined the impact of learning analytics on student outcomes based on 254 papers and found that there was limited evidence of
improvement in learning outcomes (Viberg et al., 2018). Viberg et al.’s (2018) study was broader than Bodily and Verbert’s (2017) in that they included not just student-facing dashboards but all types of learning analytics tools (i.e. included those aimed at tutors and administrators) but similarly they argue that the field needs to employ more mixed method and qualitative studies to understand how students interact with learning analytics and in order to better understand and optimise learning and the environments in which it occurs.

Both these meta studies undertaken by Bodily and Verbert (2017) and Viberg et al. (2018) point to the need to understand students’ views of dashboards qualitatively and as “tools in use” rather than as technical artefacts. This conception foregrounds the purpose of the tool in that it considers how students make use of information presented via the dashboard and how it might change their behaviour as a result. Framing dashboards as “tools in use” focusses attention on the perspective of the user, the learner and on the purpose of dashboards as tools that provide feedback to students to encourage them to make more informed decisions about their study behaviours (Bennett, 2017; Howell et al., 2018).

This paper draws on Prinsloo and Slade’s (2017a) ethical principles for the adoption of learning analytics as the interpretive lens. Prinsloo and Slade (2016) argue that in order to redress the asymmetrical power relationships associated with use of learning analytics in higher education, the ethical adoption of learning analytics needs to increase student agency and empower students as participants in their learning. In doing so adoption of learning analytics moves students from “quantified data objects to qualified and qualifying selves” (Prinsloo and Slade, 2016, p. 159). This conceptual lens presents an emergent way to understand design of student dashboards, which are frequently atheoretical (Jivet et al., 2017). It moves beyond the most dominant way that dashboards are being envisaged, as tools to make students aware of the progress, to develop their use to engage students in acting upon this feedback to improve cognitive, behavioural or emotional competencies (Jivet et al., 2017). The approach is supported by Wise (2014) who suggests agency is one of four principles for designing learning analytics interventions (alongside integration, reference frame and dialogue).
Thus, the paper takes a student-centred perspective to understanding learner-facing dashboards through seeing them as "tools in use". It aims to identify principles to inform the design of learner dashboards that will enhance student agency and empowerment.

Methodology

The study aimed to understand the ways that students interpreted the dashboard and the extent to which it might impact on their learning behaviours. Thus, it employed semi-structured interviews that enabled students’ responses and the meaning that they attach to dashboard elements to be gathered. The interviews were undertaken in two rounds, initially a self-selecting group of 10 students from a cohort of 178, but the second round was targeted at a smaller cohort and 14 out of 16 participated. Thus, the second round of interviews helps to overcome the bias inherent in self-selecting samples. Both data sets were analysed to provide a rich understanding of how students interpreted the dashboard elements. The academic range of the sample varied from 1st to 168th out of 178 in the group for the first round, and in the second from 1st to 16th in a group of 16 students, for a particular assignment presented on the dashboard. The dashboard displayed the overall degree classification that the student was on track to achieve, and this ranged from 51 per cent (low 2:2) to 74 per cent (1st) for the first round, and the range of participants in the second round was from 60 per cent (border of 2:1 and 2:2) to 76 per cent (1st). Slightly more students were doing worse in the assignment presented on the dashboard than their overall on-track score. Therefore, the sample had the potential to uncover a range of emotional responses to the assignment data, not just being pleased that this assignment was bringing their average mark up or disappointment that it was lowering their mark. The dashboard used in this study contained seven descriptive elements: Figure 1 shows four of these elements and Figure 2 illustrates the other three elements.
This was the first time that the students had seen their data presented in this way and the semi-structured interview format enabled them to ask if they wanted clarification about their interpretation of their data. The interviews lasted between 10 and 30 min (typically 15 min) and addressed three questions: What were your feelings on seeing the data? Whose responsibility do you think it is to act on this data? Would you take any action/do anything different as a result of reading your data?

The study was sensitive in nature, given its focus on students’ academic performance. BERA’s (2018) ethical principles informed the study. Participation was voluntary, and students’ identity has been anonymised through the use of pseudonyms. There was a responsibility to ensure that the students were supported during this process and this was achieved by preparing the data carefully to ensure that it was valid, and by helping students to interpret their data in a way that would encourage positive outcomes. For instance, explaining how the on-track score was calculated and how it will change according to future module results. The students were also encouraged to reflect on their progress and plan how to approach their final year of study. Ethical permission was given by the Ethics Committee of the School of Education at the case study University.

The analysis used a four-stage inductive approach outlined by Bryman (2012) in which themes were identified through interpreting the data in relation to the ways that it appeared to support students to become both agentic and empowered (Prinsloo and Slade, 2016). This inductive approach led to identification of four key principles which have been developed and refined through a reflective interpretative process. The aim was to identify principles that position students as active in the process of interpreting their dashboard data and enable students to translate their understandings into practical actions. The principles that were identified relate not only to the design of dashboards, but also to the way that the tools are used in practice reflecting the need to focus on “tools in use”. They identify the things that students particularly valued when they saw their dashboard and focus on how dashboard design might support positive impact on students’ behaviours.

The study has some methodological limitations in that it uses a small sample of final-year students from one academic discipline in one UK university. However, a strength of the study is that nearly a whole cohort was interviewed thus avoiding the bias that arises from self-selecting samples. The interviews provided a rich source of insight into students’ responses that enables the details of individual’s dispositions, experiences of study and other factors to be considered thus providing the study with depth and nuance through these qualitative interpretations. The study focusses on students’ responses to their dashboard, but the study was not able to track the impact that dashboards have on students’ actual behaviours which is an area that is worthy of further research.

**Findings: four principles for learner dashboards**

*Principle 1: customisable by the student*

Throughout the data, there were many examples of how students engaged in a personal interpretative process as they viewed their learner dashboard. Their interpretations appear to be influenced by a range of factors including the individual’s disposition and the way that they approach their studies, their graphical and data literacy and a range of contextual factors, such as the type of course that they are studying (face to face, online, exam orientated, compulsory attendance, professional components, etc.).

The need for students to be able to customise their dashboard was particularly evident when considering how criteria-referenced data are presented. RAG (Red/Amber/Green) rating is a common technique for presenting data because it conveys simply three benchmarked levels of achievement (You, 2016). In Figure 2, the Course Summary shows a RAG-rated flag to indicate a student’s performance with green for “good” which was
allocated to scores over 60 per cent, amber suggesting a result that was acceptable for scores 50–59 per cent and red for scores below 50 per cent that could suggest a problem. The data illuminated the way that RAG rating imposed the institutional perspective on the learner in the way that the flags had been colour coded according to the institutional priorities. Some students wanted to set their own RAG criteria to reflect their own personal goals (for instance Jasmin and Asmah) whilst others were happy for the institution to set the flags (for instance Marcia):

The thing about the green flag is some people will be getting a 2:2 and actually that will be an incredible grade for them. For me, I obviously want a first, and it is possibly still doable as long as I work my socks off. (Jasmin)

I don’t really get this bit, because this is green, and it’s just a B. (Asmah)

I’m happy for it [the flag colour] to be decided for me because I think it motivates me more to work harder, whereas if I set it myself, I’d just [set it] too low. (Marcia)

The student’s interpretation of the data and how it contrasts to dashboard displays that are pre-determined is clear in these quotes: both Asmah and Jasmin wanted to get a first and thus are aiming for marks in the 70s so did not like the green flag being used for marks in the 60s.

The need for customisable displays is supported by Sluijter and Otten (2017) who noted some students were satisfied with the minimum required score to pass whilst other students may aim to get the highest mark possible. As dashboards become technologies owned and supported by the institution, it is likely that their adoption will tend to be shaped by the institutional values, indeed as the institution spends large sums of money to buy proprietary systems, they will expect to see “return on their investment” driven by institutional measures of success such as retention. Placing a focus on the customisation of a dashboard helps to redress this institutional focus.

The issue of comparison with others in their cohort resulted in a range of reactions reflecting the student’s disposition and this supports the principle of allowing them to customise what they see. In Figure 1, the elements in the top and bottom right give the student information about where they sit in relation to the others in the cohort. In the quotes below Nulla found this comparison motivating whereas Justine preferred not to know this information, whilst Malcolm and Ingrid thought it was irrelevant:

With every class you’ve always got the people that are really smart and then, you know, the people you want to kind of be like. (Nulla)

I don’t think I need to know what position I am in […] then I kind of know that fourteen other people have done better than me. (Justine)

The average of everyone else, doesn’t really mean anything to me personally. As long as I’m doing what I need to do. (Malcolm)

There is no point in seeing an average of everybody’s marks, only because it doesn’t really matter what other people get because it’s only your marks that matter. (Ingrid)

Whether a student is able to see their performance compared to others in the group (i.e. norm referenced data presentation) is an unresolved and contentious topic in the learning analytics literature with one study identifying that students liked comparison with others (Konert et al., 2016) whereas another found that students preferred to see only their own data (Tabuenca et al., 2015) and others showing that it depended on the type of student and who they are being compared to (Davis et al., 2017). Similarly, this study identified a range of responses whilst many students valued being compared to others (e.g. Nulla and Ingrid); this was not universally experienced (e.g. Justine) and indicates the importance that students are students being given a choice in the way that their data are presented.
There are many aspects to this customisability: what is displayed, how it is ordered, the way that comparisons to the rest of the cohort are displayed (or not), how norm referenced and criteria referencing displays are employed, etc. However, the overriding message is that dashboards need to be customisable by the student to reflect the way that individuals interpret their data influenced by the student’s disposition in ways that are not easily predicted (Bennett, 2017). Similarly, Roberts et al. (2017) concluded that students should have the ability to decide if they want to see a dashboard, and whether comparison data are presented.

To summarise Principle 1 argues that because dashboards are individually interpreted then displays need to be customisable by the student to respond to their individual needs and in this way helps to support a student’s sense of empowerment and agency.

**Principle 2: foreground students sense making**

Principle 1 illustrated the personal interpretive process that students go through as they use their dashboards, and Principle 2 extends this idea and demonstrates how students can be supported in the interpretation process through the design features of the dashboards. There are many features that influence sense making including: the level of granularity, the level of aggregation, the form of display (pie chart, bar chart line graph or narrative, etc.) and the type of data, e.g. ipsative, criterion referenced, norm referenced. The following discussion of Principle 2 illuminates some of the features that support student agency and empowerment; however, a detailed discussion of data visualisations is beyond the scope of this paper. (For discussion of this topic see for instance Sedrakyan et al.’s (2018) mapping of data visualisations and educational conceptions.)

Students use their understanding of their particular experiences and context to inform their interpretation, so designs need to support this process. For example, when looking at the data students can draw on their memories of the experience of studying to interpret what they are seeing. Dashboards can be designed in ways that support or hinder the process of interpretation. For instance, the level of granularity in the data needs to enable a student to pinpoint precisely features of the data; being able to identify the source of the data (i.e. to distinguish between attendance from virtual learning environment, VLE clicks) helps students to interpret the data as well as to trust their validity. Hence, dashboards need to facilitate this rather than aggregating several sources together. Malcolm knew that the timing of a particular assignment was when he was dealing with some particular personal difficulties:

The reason I got this marks is due to family circumstances. I just had all on to get it in and it was just rushed. (Malcolm)

For Sarena knowing which subject the mark relates to helped her to assess its significance:

I didn’t really like that course anyway. (Sarena)

The elements were presented in a variety of different forms including the use of graphs, bar charts, pie charts and words (see Figures 1 and 2) and it was evident that they affected the interpretation process:

They [the RAG rating flags] make it more obvious, because obviously you know green’s good and red’s bad and then yellow’s like average. (Sarena)

I’m not really a bar chart kind of person. (Lydia)

Students wanted to be able to select which displays were most relevant to their course study patterns. For instance, if they were attending a course that was primarily delivered face to face they did not see any significance in the VLE data:

I don’t care how much I’ve spent on it [the VLE]. (Nulla)
Similarly, some were hostile to the display element which showed their usage of the library saying that they had other ways of engaging in their studies:

Don’t feel this is relevant as people may have bought the books. (Pavan)

It’s pointless to know how many visits you’ve made. (Rebecca)

Hence, Principle 2: foreground students sense making is significant because it draws attention to the student as the agent doing the interpretation and draws attention to the design features that support this interpretation process including the form of the display, the way that students can apply their understanding of context to understand the data given its level of aggregation and granularity.

**Principle 3: enables students to identify actionable insights**

One definition of learner analytics is that it should identify actionable insights: “Analytics is the process of developing actionable insights through problem definition and the application of statistical models and analysis against existing and/or simulated future data” (Cooper, 2012, p. 3). However, this definition gives power to the learning analytics and its algorithms. It suggests that the dashboards will provide the student with the insights. Instead Principle 3 focusses on how students can be active in the interpretation process and the ways that learner dashboards design, enable and encourage students to become active and empowered in relation to their studies. This principle could be realised in various ways, including helping the student to identify goals or behaviour change. It might also be achieved through providing displays with a level of granularity that enables the student to contextualise and interpret the display based for their course requirements and study patterns. Dashboards can support students to be active and empowered by enabling and encouraging students to interrogate their displays, to drill down and to identify how the display is constructed and how the student’s behaviours have affected the data being displayed. As suggested by Principle 2, dashboards need to foreground the student to understand the provenance of the data (rather than aggregating data), support sense making and have the potential to enable students to identify actionable insights.

The analysis demonstrated that learner dashboards were generally motivational in that many students identified actions that they would take that support positive engagement in their studies:

I need to get on those computers and check the marking criteria. (Asmah)

It shows me that I should be putting in a lot more hours than I am. (India)

Changing behaviour is often challenging, see for instance Heath and Heath (2010), and learning behaviours are similarly complex. Whilst the analysis presented here suggests that students will go on to make changes in their behaviour, for instance India suggests that she will put in time to her studies, a longitudinal approach would be needed to establish whether this change did occur. Notwithstanding, Principle 3: enable students to identify actionable insights is a central principle because it places the student as the active agent in the interpretation process. It focusses on the need to consider how students might change their behaviour as a result of using the dashboard.

**Principle 4: dashboards are embedded into educational processes**

Context plays a significant role in the uptake of all technology artefacts into educational practices, and as discussed above, these need to be viewed as “tools in use” (Boyle and Ravenscroft, 2012; Lund and Stains, 2015). Principle 4 focusses on the way that dashboards
need to be embedded into other educational processes and draws attention to the other people and systems that influence the value, relevance and significance of their adoption.

Within the institution, each student is allocated to a Personal Academic Tutor (PAT), who is expected to meet with their tutees five times a year. These meetings provide students with someone who knows about their academic development and could support students in the interpretation of their dashboards and help them to identify goals and actions. The dashboard appeared to offer potential to enhance the PAT–student relationship, leading to more informed conversations about a student’s progress and other development and support needs:

So if I can provide this data to them [their PAT] and they can help me to push to a first, I think it’s really helpful. (Harry)

I think it [taking the data to the PAT] would be very handy […] because then they’d be able to assess your data and you’d be able to discuss where you’re falling short and where you’d be. (Lydia)

In addition, the students felt that having the dashboard data would give their PATs a more well-rounded view of their engagement and effort beyond just their assessment marks:

For even the tutors to see the effort you’re putting in. (Lydia)

It needs to be recognised however that dashboards are a part of a much wider system of feedback and support for students and therefore need to be considered within this context. They are just one tool an institution has at its disposal to help support students, but at the heart of the student experience is human interaction. Indeed, others have identified the significance of human mediation in relation to student support (Thompson and Mazer, 2009) and as illustrated above, the dispositional aspects of interpreting dashboards appear significant. Torrance (2012) reminds us that “providing and receiving feedback is a highly demanding emotional process, impacting on learners’ identities and notions of self-worth” (p. 334).

The fourth principle for dashboard design is therefore to make sure that the student dashboard is embedded into other educational processes these might be a formal academic support system, or a reflective developmental process and that include human interaction.

Conclusions

There is a growing use of data within higher education, yet to date, most of institutional adoption focusses on how this can be used to support the organisation’s processes and practices, with little attention being paid to students’ perspectives and how data might support the student experience (Roberts et al., 2017). This study has addressed this gap by understanding students’ responses to data. In doing so it takes up the moral imperative articulated by Slade and Prinsloo (2013) that learning analytics should work with students so that they are “engaged as collaborators and not as mere recipients of interventions and services” (p. 1519). This paper illustrates students’ responses to receiving data about their learning presented to them in dashboard format and draws on an interpretive lens of student agency and empowerment (Slade and Prinsloo, 2013) to identify four principles to inform the design of learner dashboards. The four principles, designs that are customisable by students; foreground students sense making; enable students to identify actionable insights; and dashboard use is embedded into educational processes, provide a road map for developers and higher educational institutions in guiding their design and implementation of learner dashboards. The significance of these four principles is that they shift our attention from the technical features of dashboard design to foreground instead the ways that they are interpreted by students. In doing so the principles emphasise values of student agency and empowerment which Prinsloo and Slade (2016) propose as a key tenet of ethical adoption of learning analytics.
The principles have significance particularly for developers of dashboards to guide and evaluate their designs. For instance, Principle 1: customisable by the student suggests practical ways that the function and form of dashboards might be designed so that students can tailor the displays to suit their particular needs which would reflect their aims and aspirations and their personal dispositions (e.g. response to peer comparison). Principle 2: foregrounds students’ sense making suggests making available the data sources so that students are able to interrogate these, and not aggregating sources so that their provenance is lost. It suggests enabling students to contextualise the displays to their particular programme and circumstances. Principle 3: enables students to identify actionable insights suggests that designs should help students to interpret the data rather than digesting them for the student. This might mean providing guidance to the student about what is on view and prompts to help them to identify what actions they could take next.

The final principle, dashboards are embedded in educational processes, places attention on the ways that the tools are integrated into the educational systems and processes. We know that uptake of learner dashboards has been low (Bodily and Verbert, 2017) so a key focus in terms of the design is how it is embedded into student behaviours. The literature about how students respond to feedback draws our attention to the importance of moving from “transmission of comments from marker to student, towards a more dialogic focus on student engagement and the impact of feedback on student learning” (Winstone and Boud, 2018, p. 1) and emphasises the role that staff play in facilitating students’ understanding of feedback (Carless and Boud, 2018). Similarly, for the learner dashboard to have an impact we need to ensure that institutions embed this new technological tool in teaching and learning regimes and draw on lessons of technological adoption (see e.g. Brown, 2013).

References


25

Student agency and empowerment


**Corresponding author**
Liz Bennett can be contacted at: e.bennett@hud.ac.uk
Leveraging learning analytics for student reflection and course evaluation

Devrim Ozdemir
College of Health Sciences, Des Moines University, Des Moines, Iowa, USA, and
Heather M. Opseth and Holland Taylor
Physician Assistant Program, Des Moines University, Des Moines, Iowa, USA

Abstract

Purpose – The purpose of this paper is to demonstrate a process of faculty utilization of learning analytics by evaluating students’ course objective achievement results to enable student reflection, student remediation and faculty curriculum evaluation.

Design/methodology/approach – Upon the completion of a backward curriculum design process, the authors utilized learning analytics to improve advising, student reflection, remediation and curriculum evaluation. The learning management system incorporated the learning analytics tool to assist the learning analytics process. The course faculty, student advisors and students utilized the learning analytics throughout the academic year.

Findings – Unlike relying merely on student grades and other proxy indicators of learning, the learning analytics tool provided immediate and direct data to multiple stakeholders for advising, student reflection, student remediation and course curriculum evaluation. The authors believe it was a meaningful endeavor. It enabled meaningful conversations focusing on course learning objectives and provided detailed information on each student. The learning analytics tool also provided detailed information regarding which areas faculty needed to improve in the curriculum.

Originality/value – Most of the literature on learning analytics present the cases that administrators utilized learning analytics to make higher level decisions and researchers to explore the factors involved in learning. This paper provides cases to faculty regarding how learning analytics can benefit the faculty and the students.

Keywords Assessment, Learning analytics

Paper type Case study

1. Introduction

As the learning analytics concept continues to gain more interest in higher education, the utilization of learning analytics to improve student learning becomes more critical for student success. The literature sheds light on the many uses of learning analytics mostly to collect and analyze student achievement data in an aggregate fashion (Abdous et al., 2012; Falakmasir and Habibi, 2010). Higher education administrators are also interested in learning analytics to comply with internal and external regulations (Dziuban et al., 2012).

In this paper, the authors provide a case study in which they utilize learning analytics at the course level to provide direct feedback to the students for reflection and remediation purposes in a graduate-level health sciences program. The learning analytics reports are also utilized to evaluate the course curriculum to improve areas which may eventually increase student learning performance. To take advantage of learning analytics, the authors followed a granular curriculum alignment process over the last five years. During the alignment process, faculty developed measurable and specific course learning objectives and aligned them with program-level competencies. Those objectives are also aligned with summative assessments to measure each course learning objective directly. Authors consider summative assessments as graded assessments which may include multiple-choice and rubric-based assessments. The learning analytics system allows students to view their course learning objective achievement progress throughout each course along with other
student performance data such as course login data, content visit statistics, online discussion participation and grades. When course faculty observes significant student struggle in the course, the faculty requires students to view course learning objective achievement for reflection and remediation. Course faculty determines the need for reflection and remediation using learning analytics tool and, in particular, to evaluate the individual student’s achievement of the course learning objectives. Along with course faculty, student advisors also view student achievement to monitor the student academic progress and to meet students to discuss their progress. Course faculty utilizes the system to evaluate the course curriculum and to identify course deficiencies to drive improvement at the end of the course. Overall, comprehensive learning analytics data allow stakeholders (course faculty, student advisors and students) to recognize program strengths and weaknesses.

In the following sections, this paper presents a literature review on learning analytics, the approach taken by the authors to utilize learning analytics, the findings and the conclusion.

2. Literature review

According to a recent literature review on learning analytics in higher education, 1,434 papers have been published between the years of 2012 and 2018 (Viberg et al., 2018). This number demonstrates the increasing popularity of learning analytics in higher education. Learning analytics became popular for a variety of reasons. Higher education institutions have experienced increasing scrutiny over the years with the increasing demand for accountability from the public and profession (Goldstein and Katz, 2005). Learning technologies such as learning management systems have improved over the years and developed learning analytics tools which were not available in the past. The data in those systems continue to be more accessible to the end users. In the scholarship of teaching and learning, learning analytics along with educational data mining and academic analytics continue to receive more attention which makes learning analytics research more appealing.

Pioneer instructors continue to search for innovative strategies to gain a better understanding of what and how students learn in their courses. When all these factors combine, stakeholders from different fields continue to benefit from learning analytics. Learning analytics provide direct and indirect data and opportunities to analyze students’ learning progress and performance. As a result of a wide variety of stakeholders, the definitions and uses of learning analytics vary.

Long and Siemens (2011, p. 32) stated that “higher education, a field that gathers an astonishing array of data about its ‘customers,’ has traditionally been inefficient in its data use, often operating with substantial delays in analyzing readily evident data and feedback.” These data have been collected in many electronic platforms as well as on paper. From electronic student information systems and learning management systems to other academic units providing curricular, co-curricular and extra-curricular programs and services, many stakeholders in higher education institutions collect various kinds of student data. Depending on the stakeholder collecting and analyzing the data, the processes could take different names or perhaps use the same names with a completely different meaning behind them.

There are three familiar concepts relevant to the analysis of student data: educational data mining, academic analytics and learning analytics. Arguably, the most significant difference between educational data mining and the analytics approaches is if a question or an apparent problem initiates the learning analytics. Baepler and Murdoch (2010, p. 2) explained that “analytics is associated with a scientific, hypothesis-driven approach, while data mining has a legacy with strategic business techniques and marketing.” While researchers and practitioners could use either approach, the main difference between the two approaches is analytics start with a pre-established notion while educational data mining attempts to learn from the data without any pre-established notion (Baepler and Murdoch, 2010).
Therefore, analytics approach attempts to contribute to an existing theory or to address important issues such as the most significant independent factors associated with student dropout rates (Morris et al., 2005) or student learning in a course (Abdous and Yen, 2010).

Recent literature on educational data mining revealed 1,629 studies published between the years of 1983 and 2016 (Dutt et al., 2017). Educational data mining is a growing area of interest mainly for researchers with information technology and computer science background. Due to its complexity, however, educational data mining may not be a feasible option for faculty and students to have direct benefits. Perhaps, it has not been the primary goal of educational data mining to become a tool for the direct use for faculty and students. It is important to note that educational data mining studies can be critical to understanding specific factors involved in a classroom environment which previously known theory-based studies failed to address. However, it is also important to note that depending on the type of data involved during the educational data mining process the results may not be generalizable into different settings. For instance, Falakmasir and Habibi (2010) found that virtual synchronous classroom participation has the most impact on students’ final grades compared to other factors such as asynchronous discussion forum participation and access to archived materials in a virtual course. However, another study using educational data mining failed to demonstrate the relationship between students’ number of questions, chat messages, login times and student success (Abdous et al., 2012). The comprehensive critical analysis of educational data mining against other approaches is beyond the scope of this paper.

Educational data mining may use the same data but may apply unique data modeling techniques to explore what the existing data demonstrate rather than asking any specific questions. Educational data mining follows four common steps: collect data; preprocess data; apply data mining; and interpret, evaluate and deploy the results (Romero et al., 2008). The educational data mining techniques include neural networks, genetic algorithms, clustering and visualization methods, fuzzy logic, intelligent agents and inductive reasoning (Castro et al., 2007). Educational data mining studies demonstrated significant findings which contributed to the understanding of learning particularly in online environments (Abdous et al., 2012; Falakmasir and Habibi, 2010).

Along with educational data mining, academic analytics and learning analytics are also familiar concepts in the field. The most significant difference between academic analytics and learning analytics is the type of questions stakeholders ask and the beneficiaries of the analytics. While learners and faculty benefit from learning analytics, administrators, funders, marketing, national governments and education authorities benefit from academic analytics (Long and Siemens, 2011). Baepler and Murdoch (2010, p. 1) refer to academic analytics as “a data-driven decision making practices for operational purposes at the university or college level, but it can also be applied to student teaching and learning issues.” A survey study revealed that stakeholders utilized academic analytics in advancement/fundraising, business and finance, budget and planning, institutional research, human resources, research administration and academic affairs (Goldstein and Katz, 2005). However, challenges and risks related to the academic analytics such as “big brother,” the possibility of error, obligation to act, distribution of resources and profiling were reported (Campbell and Oblinger, 2007).

In contrast, learning analytics is defined as “the measurement, collection, analysis, and reporting data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs” (Siemens, 2011). In that regard, researchers investigate direct and indirect factors that may impact learning in various environments by using learning analytics tools. For instance, Lu et al. (2018) investigated 21 variables on student final academic performance such as video lecture analytics, online course login data, and student assessment scores in hybrid learning environments. Another similar study found that learning analytics data such as login
engagement, content visit statistics, formative assessment results, posting engagement in the online discussions could predict the final score students receive with 63.5 percent accuracy (Saqr et al., 2017). Many studies approached the learning outcome as the final grade received in courses. However, metrics collected from learning management systems could also be proxies rather than the accurate indication of learning (Dietz-Uhler and Hurn, 2013).

Unlike the majority of the learning analytics focusing on the investigation of the factors affecting learning, this paper focuses on providing actual practical cases where course faculty members utilize learning analytics to assist students in the reflection of their learning and evaluate course curriculum. The utilization of learning analytics by the faculty and student through pedagogical interventions is particularly important for student success (Wise, 2014). Unlike previous studies, the selected metric for learning in this paper is the achievement of course learning objectives rather than final grades in the courses. This way, students would be able to receive direct feedback on the course learning objectives as they expect to receive from learning analytics systems (Schumacher and Ifenthaler, 2018). Also, the authors use learning analytics to evaluate the curriculum based on the educational interventions provided to the students. These educational interventions are critical for quality education (Snodgrass Rangel et al., 2015).

3. Approach
This paper depicts the process of utilizing learning analytics for advising, remediation, student reflection and curriculum evaluation. Authors describe learning analytics as a process of assessing students’ learning by examining their achievement on course learning objectives with the existing student performance analysis tools embedded into the learning management system. In order to benefit from the tool, the degree program has gone through a comprehensive curricular alignment process. This alignment process allows the instructors and students to precisely monitor the progress of each course learning objective. The learning analytics process allows the instructors to enable student reflection during one-on-one academic advising sessions and remediation sessions as well as to evaluate the course curriculum for further improvements at the end of the course.

3.1 Context and demographics
The degree program is a face-to-face 25-month long Master of Science in Physician Assistant Studies program and is housed in a graduate health sciences university in the Midwest, USA. The program has been training students since 1981 and has maintained its special accreditation since then. Graduates of the program gather patient histories, perform physical exams, order and interpret diagnostic studies, formulate diagnoses and develop pharmacologic and non-pharmacologic management plans under the supervision of a physician. The program accepts one cohort at the beginning of each academic year with approximately 50 students per cohort. Students entering the program are adults and predominately female. They have earned a bachelor’s degree from a regionally accredited institution in the USA. In the first (didactic) year of the program, students develop the necessary clinical knowledge, skills and attitudes necessary to enter into clinical settings. In this didactic period, the program offers clinical courses comprised of several subsections each representing an individual organ system in addition to basic science, physical diagnosis and clinical assessment courses. In the second (clinical) year, students expand their knowledge and practice their clinical skills through supervised clinical practice experiences. The program assesses students through a variety of methods, such as multiple-choice exams, rubric-graded projects and direct observations. Faculty conducts formative and summative assessments of student learning throughout the program to assess student learning in each course and the entire program.
3.2 Curriculum alignment process
In order to obtain the direct data on course learning objective achievement, the program has gone through a backward design process (Wiggins and McTighe, 1998) in which faculty have aligned program-level learning outcomes with individual course learning objectives. Faculty also aligned the same objectives with summative student learning assessment sections. Figure 1 illustrates the curriculum alignment process.

3.3 Learning analytics
Desire2Learn Brightspace learning management system and Desire2Learn Insights tool have been used to collect and report learning analytics data. Faculty used learning analytics for the summative assessments of student learning at the individual course and program-level. Faculty considered summative assessments are as those comprehensive graded assessments. The summative assessments directly assess the course learning objectives. Learning analytics provide both direct and indirect data reflecting on the quality of student learning. Indirect data include student course login information, the amount of time spent on teaching and learning activities, and participation in asynchronous online discussions. Direct data include student performance on individual learning objectives at the course level and program-defined learning outcomes at the program level. Therefore, direct data include grades received for sections of summative learning assessments. Either a specific question pool in a multiple-choice exam or a specific criterion in a grading rubric aligned with an individual course learning objective represents a section. This process allows the program to evaluate student achievement at a granular level. Each course learning objective assessment receives a corresponding threshold in which the individual student must meet to obtain a checkmark in the dashboard denoting achievement of the corresponding learning objective. If one fails to meet the assigned threshold, learning analytics tool shows a failure to achieve with a cross sign (X). Figure 2 below illustrates the student view of the achievement results of the course learning objectives for one example student in a selected course.

3.4 Application of learning analytics
During the implementation of a course in the program, the instructors utilize the learning analytics for three specific purposes (academic advising, remediation and curriculum evaluation) and the students use the analytics for self-reflection. Throughout the academic semester, the instructors apply the learning analytics for academic advising and students apply the learning analytics for self-reflection. At the end of each summative assessment of student learning, the instructors review the individual student summative learning assessment and schedule a one-on-one meeting with those who are expected to reflect on their areas for improvement based on the learning analytics results.

Additionally, the students who fail to pass a subsection of the course are expected to complete remediation to address the areas for improvement depicted in the learning analytics system. The same results are also utilized in an aggregated manner to evaluate the course curriculum as a whole at the end of the course to make the necessary curricular improvements for the upcoming course offerings. The following sections provide cases in

Figure 1. Depiction of the curriculum alignment
which learning analytics have mainly been used to assist student self-reflection, as well as the instructors in the utilization of learning analytics to evaluate the course curriculum.

3.4.1 Advising. Learning analytics provide student advisors with valuable information about individual student performance at the course level. The program assigns academic advisors to students after their admission. Student advisors meet with students regularly to discuss academic progress on assessments throughout the curriculum. The advisors leverage learning analytics to identify student areas for improvement at the level of the individual course learning objective. Advisors can identify trends in areas of struggle for students which are of interest to students and their advisors. Faculty and advisors tailor advice on how to improve knowledge on course content by developing individualized study plans or recommending resources to improve the deficient area. For example, if a student has not met benchmarks for instructional objectives across multiple courses around a specific topic, the student is directed to resources to improve specific knowledge in the area.

3.4.2 Student remediation. In addition to individual advising purposes, course instructors and students use learning analytics to identify areas for which students need remediation at the course level and develop remediation assignments for each student’s needs. In longer courses which span an entire academic year and have many assessments focusing on distinct areas such as individual organ systems, students may be able to pass the entire course while struggling in a particular organ system area. Since each organ system has equal significance and students must separately master, the program developed a remediation process within individual courses to ensure that students demonstrate sufficient knowledge on each organ system regardless of what their final course score is. For example, a large clinical medicine course consists of hematology, respiratory, gastrointestinal and cardiovascular systems. If the students fail to obtain the threshold (75 percent) for any individual system, they are required to complete a remediation assignment before being allowed to advance through the curriculum. The remediation assignment expects each student to address the failed course systems.
learning objectives in the particular organ system. The course instructor directs the student to review the learning analytics data and requires students to complete the remediation assignment. The students must demonstrate the sufficient understanding relevant to the content covered by the previously failed instructional objective by creating an essay supported by external references.

3.4.3 Student reflection on overall learning. One example of the student self-reflection occurs at the end of the program within the last course. Students complete a comprehensive final exam, which mimics the national certification exam. In the last month of the program, the program assesses students on overall knowledge gained throughout the didactic and clinical curricula with a multiple-choice comprehensive final exam. Upon completion, students can reflect on their learning performance on 13 organ systems and seven task areas within each system. The learning analytics tool provides a detailed report to students for reflection. This reflection enables students to study for the national certification exam appropriately.

3.4.4 Curriculum evaluation. After completion of a course offering, course instructors evaluate their courses through a data triangulation process. The data come in both quantitative and qualitative format. Faculty uses this data to develop a comprehensive evaluation report. This report includes student evaluations of the course and instructors as well as aggregate course learning objective achievement results. The program requires all students to complete course and instructor evaluations. Faculty evaluates course learning outcome achievement on all instructional objectives by generating a “Course Learning Outcome Evaluation” report in the learning management system as depicted in Figure 3. The report provides a list of all course objectives with corresponding overall student achievement percentages. If the course learning objective achievement percentage rate is below the programmatic benchmark, the program needs to investigate further to identify the potential causes. The reasons for lower percentages may be relevant to instructors’ teaching styles, delivery of course content, assessment structure, alignment between the course learning objectives and assessments, the sufficiency of student preparation or the difficulty level of the content. Faculty conducts further investigation to identify the root cause of the problems when needed.

4. Findings
In this section, the paper presents the findings of the case study in the specific areas of advising, student remediation, student reflection and curriculum evaluation. The authors identified improvement areas and developed realistic and timely plans for improvement for

<table>
<thead>
<tr>
<th>Course Learning Outcomes: 8</th>
<th>Total Number of Participants: 51</th>
<th>Number of Assessment Activities: 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Outcome:</td>
<td>Achievement Results</td>
<td></td>
</tr>
<tr>
<td>NUT1: to recognize key trends in nutritional epidemiology and identify how dietary changes and the focus on prevention have impacted human health and disease</td>
<td><img src="image1" alt="Achievement Results" /></td>
<td></td>
</tr>
<tr>
<td>NUT2: to identify the role of dietary carbohydrates and their relationship to disease</td>
<td><img src="image2" alt="Achievement Results" /></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3. Introduction to Nutrition sample course learning outcome evaluation report

Leveraging learning analytics
students by utilizing learning analytics tools.Below are the findings based on the authors’
experiences during the utilization of learning analytics for advising, student remediation,
student reflection and curriculum evaluation.

4.1 Instructional and technology training is required for the stakeholders
In order to utilize learning analytics, authors needed to have access to students’
learning analytics data as well as to possess the intimate knowledge of technology.
The authors observed that the advisors needed to be familiar with the structure of the
courses to provide the advice. At times, the authors struggled to provide the advice
when data came from multiple courses during their sessions with the students. The
difficulty of integrating learning analytics into the regular advising practices also was
addressed in the literature (Krumm et al., 2014). The authors heavily relied on their
technical knowledge to benefit from learning analytics. Therefore, the institution should
provide training for educational stakeholders on using learning analytics in order to be
successful (Scheffel et al., 2014).

4.2 Additional time allocation needs to be provided to faculty and advisors
Utilization of learning analytics to focus on the students’ course learning objective
achievement has been beneficial for the faculty to accomplish more evidence-based advising
and remediation while allowing the student to reflect on their learning. Learning analytics
also allows the faculty to conduct more meaningful curriculum evaluations by focusing on
student learning performances. In order to obtain the course learning objective achievement
results, the faculty, however, needed to invest additional time to develop and align the
course objectives properly to summative assessments. The time investment continues to
exist for faculty to maintain the alignment. There have been many cases in which the faculty
decided to update the course learning objectives upon their evaluation. They needed to work
with the instructional design coordinator in their college to maintain the curricular
alignment of their courses. The change in the objectives creates a trickle-down effect due to
the alignment with their alignment with program-level expectations, summative learning
assessments and the instructional activities. The maintenance of the alignment also
becomes more demanding in comprehensive exams due to a considerable number of
connections that faculty and the instructional design coordinator need to make at the back
end of the technology to feed the learning analytics system.

The learning analytics tool provides a large quantity of data. Programs must determine
at which checkpoints throughout the curriculum it is important to assess students
in each of the competencies, and what level of performance on all competencies
students must demonstrate at each checkpoint. The program may identify some
competencies that require a higher level of achievement than others, depending on the
specific learning outcome identified by the program. Colvin et al. (2017) also emphasizes
how staff and institutional capacity could be critical for the institutional adoption of
learning analytics.

4.3 Early introduction of learning analytics to the students is critical
The program needs to introduce students to learning analytics early and often in the
program to maximize its benefits. Wise (2014) describes this process as “grounding” under
the “integration” principle for the “pedagogical learning analytics intervention design”.
According to Wise (2014), the grounding process intends to help the students to see the
connections between learning design and the use of learning analytics. The degree program
perhaps could deliver such an introduction during the student orientation. Otherwise,
authors observed that it took additional valuable faculty and advisor time to introduce the
tool later when needed rather than focusing on what the tool provided. In addition to the introduction to the learning analytics tools, the degree program should present students the value of self-reflection on their learning using the learning analytics.

4.4 Learning analytics data could be informative for both students and faculty
Long and Siemens (2011) explains learning analytics benefit various stakeholders in education. Authors experienced learning analytics data as informative to both students and faculty. Authors observed that students received immediate feedback on their course objective achievement with the learning analytics tool. This informative aspect of learning analytics is in line with students’ expectations on personalized analyses of their own learning activities (Schumacher and Ifenthaler, 2018). The authors also had the opportunity to evaluate the course curriculum based on the learning analytics data. The structure and timing of the program’s assessment alignment allow faculty to assess performance trends over time and the data aids in the identification of students who require remediation. While standardized assessments are great for comparison to national means, home-grown exams allow for greater flexibility, as the faculty aligns each assessment item to instructional objectives, which further align with competencies for the profession. Pulling data from an individual competency allows the faculty to assess students from a different perspective. Rather than assessing competency through points on an exam, or percentage in a course, faculty could look at overall performance related to the competencies deemed necessary for entry into clinical practice. This level of assessment for proficiency in each competency domain is critical in medical education, due to the level of knowledge/skill/ competency expected of a graduate upon entry in clinical practice.

From a technical standpoint, the learning management system followed an “all” criterion for objectives. This particular criterion created a limitation regarding how the objectives could align with the assessments. If an objective aligned with multiple assessments, the student was expected to pass the threshold for all aligned assessments. The limitation of the learning analytics system provided “false negatives” in which student demonstrated the achievement of the course learning objective at least once but received a failing status due to failing to achieve all corresponding assessments.

5. Conclusion
This case study concludes that learning analytics could be promising for student reflection, student remediation and faculty course curriculum evaluation. However, findings of this study provided necessary cautions for future researchers and practitioners. Moving forward, researchers may consider the investigation student and faculty self-perceptions during the leveraging of learning analytics for student reflection and course evaluation. Considering this study’s limitations, the results of a bigger scale study with the direct survey of students and faculty may shed light on student and faculty attitudes toward the utilization of learning analytics. In addition to survey studies, focus group interview studies may bring more depth as a follow up to this study.

Another important future research could be on student and faculty behaviors during the use of learning analytics. This future study may be critical to develop the grounding process during the integration of learning analytics as Wise (2014) mentioned. Although this study focused on the utilization of learning analytics from a practical perspective in a graduate health sciences program, the results may differ in other fields. Authors recommend future replication studies to identify if the findings would be similar in other fields.

From a practical standpoint, the authors found learning analytics very beneficial for advising, remediation, student self-reflection and curriculum evaluation. Therefore, the authors recommend other practitioners to consider learning analytics for similar purposes. As higher education institutions continue to gain more interest in the utilization of the
learning analytics, it is expected that more benefits of learning analytics and best practices will arise for practitioners. Learning analytics would provide immediate data on student achievement and curriculum performance. The stakeholders would be able to provide the necessary interventions more efficiently. Learning analytics would equip and empower both students and faculty during their journey of learning and teaching.

The authors will continue to use student performance data provided by the learning analytics tool along with the course learning objective data. Such data include login information, time spent on a particular recorded video lecture or stand-alone reading materials. The program will also continue revising the advising, remediation and curriculum evaluation processes to optimize the workload both for faculty and students. The plans also include the utilization of a business intelligence system such as Microsoft Power BI to be able to triangulate data coming from different systems more efficiently. Such data includes student course evaluations, preceptor evaluations during the clinical year, as well as student performance data including course learning objective achievement data. Moving forward, the authors anticipate the curriculum evaluation data will become very critical during the continuous quality improvement efforts.

References


About the authors
Devrim Ozdemir, PhD, is an award-winning accomplished scholar with three articles and one book chapter published in the area of competency-based education in the past three years. The co-author worked with three graduate health sciences programs to align their curricula. This process included working with more than 60 faculty members and over 80 required and elective courses. The co-author presented the results of the competency-based curriculum alignment initiatives to both internal and external stakeholders and assisted programs with accreditation. The co-author published articles in health administration, public health and physician assistant education journals, as well as a book chapter entitled, A Framework for the Evaluation of Competency-based Curriculum. The co-author was recognized for his work and earned an innovator’s award at state level and current university’s outstanding service award. Devrim Ozdemir is the corresponding author and can be contacted at: devrim.ozdemir@dmu.edu

Heather M. Opseth, MPAS, PA-C, has professional experience in an academic setting as an academic coordinator and assistant professor. Duties of the position include course coordination which entails instructional objective writing, alignment of objectives to programmatic competencies and course analysis. Within the PA program, she works closely with other faculty and instructional designer to maintain the programmatic alignment to maintain accreditation.

Holland Taylor, MSPAS, PA-C, has three years of professional experience in curriculum alignment through the coordination of over 20 course offerings. These duties include instructional objective writing/revision, linking to PA competencies, creating effective aligned assessments and working with the Instruction Design Coordinator to build the model within the learning management system. As Assistant Director, part of the author’s role is to manage the PA program’s curriculum and communicate details of the curriculum to other programs and departments on campus. The author has presented our assessed competency-based alignment process and associated generated data analysis internally to students, faculty and staff, and externally with accreditation representatives.

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Engaging online students through peer-comparison progress dashboards

Paula Smith
Department of Clinical Surgery, University of Edinburgh, Edinburgh, UK

Abstract

Purpose – Students studying exclusively online face the challenge of gauging their progress in relation to that of their disparate peers. The purpose of this paper is to describe the creation of a student progress “dashboard” in an online Masters programme, and the perceived effectiveness of the tool for engaging students.

Design/methodology/approach – Tableau® visualisation software was used to create a dashboard displaying cohort comparison data comprising metrics relating to the continuous assessment components of the Masters programme. An anonymous questionnaire gauged students’ perceptions of the dashboard.

Findings – Feedback from students (n = 137) suggests the dashboard improved their motivation, incentivising change in study behaviours, and sense of belonging to an online community of learners. It also acted as a conversation catalyst between staff and students, whereby students more readily engaged in dialogue with their personal tutor.

Practical implications – Distance learners are more likely to feel isolated and can become demotivated, which contributes to typically higher levels of withdrawal from online programmes vs those delivered on-campus. Tutors may consider communicating progress data as dashboards to enable online students to monitor their academic progress alongside that of their peers, as a motivational tool in an otherwise disparate group of learners, and to reduce feelings of isolation by reminding distance learners that they are part of a larger online community.

Originality/value – This paper shares student and tutor perspectives on the use of dashboards to increase online students’ motivation, and examines whether the benefits of a peer-comparison dashboard are reserved for high-achieving students.

Keywords Dashboard, Online students, Learning analytics visualizations

Paper type Case study

Introduction

Successful personal tutoring in higher education, in terms of student retention and performance, includes proactive student academic and pastoral support and monitoring participation (Thomas et al., 2017; Walker, 2018). With increasing demands on academics’ time, personal tutoring has been seen as an unsustainable aspect of the role (Macfarlane, 2011). Furthermore, online, distance-learning courses present a particular challenge to tutors in enabling students to gauge their progress alongside that of their geographically disparate peers. Measures to make personal tutor interaction with online students more effective and efficient are warranted.

The purpose of this case study is to illustrate how a progress “dashboard” can be used by personal tutors for online courses to provide effective student support as well as reinforce in students their sense of belonging to an online community. The objectives of the dashboard were to permit students to monitor their progress easily in relation to their peers, and for tutors to better identify student engagement and prompt effective interventions. Students’ perceived self-efficacy affects their motivation level (Bandura, 1994); if they believe they possess the skills to succeed, they persevere regardless of challenges and can apply skills learned more effectively in the future. Such self-efficacy can be strengthened through insightful feedback on the dashboard, and is grateful to Michael Begg and Matthew Hammond (creators of the heat-map analytics service), and Wilma Alexander, all from the Information Services Group at Edinburgh University, for their helpful discussions on learning analytics.

The author would like to thank the surgical trainees who participated in the student surveys for their insightful feedback on the dashboard, and is grateful to Michael Begg and Matthew Hammond (creators of the heat-map analytics service), and Wilma Alexander, all from the Information Services Group at Edinburgh University, for their helpful discussions on learning analytics.
through social models, as Bandura observes, “Seeing people similar to oneself succeed by sustained effort raises observers’ beliefs that they too possess the capabilities to master comparable activities required to succeed.” (Bandura, 1994).

The questions addressed in this paper are:

- Do peer-comparison dashboards increase online student motivation?
- Do high-performing students benefit most from a peer-comparison dashboard?

Literature review

Online communities

Distance learners are more likely to feel isolated and can become demotivated, which contributes to typically higher levels of withdrawal from online programmes vs those delivered on-campus (Tyler-Smith, 2006); for example, Patterson and McFadden (2009) observed drop-out rates for online masters programmes exceeding 40 per cent. Student engagement and participation have to be tracked and acknowledged in order to maintain learner motivation (Lomis et al., 2017; Teasley and Whitmer, 2017), and successful online programmes typically include sustained tutor–student contact as well as student-student interaction. Discussion boards create a collaborative online community that can reduce feelings of isolation associated with distance courses (Mohamad and Shaharuddin, 2014; Uijl et al., 2017), and contribute to increased learning and student satisfaction (Rovai, 2002; Croft et al., 2010). Furthermore, asynchronous discussion boards promote greater higher-order learning than face-to-face dialogue since students have more time to reflect upon and research their responses (Brierton et al., 2016). However, it is possible for some students to adopt the role of “lurker” (Beaudoin, 2002), in that they may view the discussions posted by others but do not themselves participate. Sharing comparative progress data among students might act as an incentive for lurkers to contribute to online discussions. For example, Bratitsis and Dimitracopoulou (2009) observed that use of interaction analysis tools which show individual student activity, as well as overall class activity, enhanced participation on discussion boards, both quantitatively and qualitatively.

Dashboards

The concept of using learning analytics to personalise each student’s academic experience is becoming a widespread phenomenon across the higher education sector (Gašević et al., 2015; Sclater et al., 2016). At the outset of this case study, in 2014, learning analytics dashboards were focused predominantly on instructor-facing solutions. The past decade has seen an expanding use of student-facing learning analytics dashboards, which facilitate student autonomy and enhance motivation beyond reporting systems used by teaching or administrative staff (see review by Bodily and Verbert, 2017a). An early adopter of these was Purdue University in the States which, in 2007, piloted Course Signals, a traffic light system to show how students were performing on their courses, which was automated and rolled out across the university in 2009. This simple visualisation tool is followed up with messages from staff suggesting what students need to do to maintain/improve their results. Data from Purdue showed that retention rates were improved and the majority of students increased their grades (Arnold and Pistilli, 2012). The emphasis of early dashboards was on retention rates and performance improvements in students at risk of failing, and many practitioners in the learning analytics field routinely analysed patterns in educational data to create algorithmic models to make predictions on academic performance (Papamitsiou and Economides, 2014). However, there are dangers in pigeon-holing any student to a predicted trajectory of failure, not least the potential for self-fulfilling prophecy (Dietz-Uhler and Hurn, 2013). Attention has turned to how student-facing dashboards can be
used to enhance student learning and reflection. To inform the design of dashboards, we need to understand how students interpret the presentation of data. A study by Corrin and de Barba (2015) revealed that while the majority of students were able to interpret the data shown and regarded them as motivational, some struggled to interpret the dashboard in a way that would inform their subsequent learning strategies. As such, the authors recommend that support is provided to students to help them interpret dashboards (Corrin and de Barba, 2015). More recently, Kitto and colleagues provided a persuasive narrative on how to design learner-centred dashboards (Kitto et al., 2017). Their “do-analyse-change-reflect” approach comprises four phases of learning analytics, the first of which involves students participating in a learning activity; second, an analysis of dashboards relating to the “do” phase; third, encouraging students to change their behaviour in light of the dashboard data; and finally asking students to reflect on the process. Kitto et al. (2017) conclude that dashboards should be integrated into the pedagogical structure of a course, coupled with assessment, to encourage their use in helping students understand and apply the data to achieve their learning goals. It is clear from the above examples that students must be at the centre of dashboard development; any data portrayed within a dashboard must not be open to misinterpretation and staff support must accompany its release.

Social learning theory and learner motivation
Observing the behaviour of others and its consequences plays a large part in learning. Bandura’s Social Learning Theory describes observational learning as “learning by example”, influenced by reinforcement whereby positive incentives can enable action by the learner (Bandura, 1971). Dashboards can be designed with a social reference frame to allow students to compare their progress with peers and modify their behaviour and performance. An example of this is increased learner engagement and completion rate in four MOOCs when a social comparison feedback dashboard was utilised (Davis et al., 2017). Students who are underperforming are the ones to derive most benefit from dashboards, but rank-order data have the potential to provide a negative incentive, and ultimately discourage lower performing students (Cherry and Ellis, 2005). Gašević et al. (2015) comment that the negative impact of comparison dashboards on students with low levels of self-efficacy “is a hypothesis commonly heard in the discussions within the learning analytics community”. However, there is a growing literature that supports the motivational impact of dashboards (Fritz, 2011; Arnold and Pistilli, 2012; Park and Jo, 2015; Bennett, 2018). Tan et al. (2016) reported mixed motivational outcomes in students who were performing below the class average; for some the dashboard stimulated competition through “healthy peer pressure”, but for others these data were “demoralising”. A recent study by Teasley and Whitmer (2017) demonstrated that students with a low grade point average (GPA) found the dashboard of more value than their high GPA-scoring peers. Furthermore, the low GPA students reported an increase in motivation after seeing the dashboard (Teasley and Whitmer, 2017). Thus, peer-comparison dashboards have the potential to prompt students at risk of failing into putting more efforts into their study but is by no means a certainty.

This original paper shares student and tutor perspectives on the use of dashboards to increase online students’ motivation, and examines whether the benefits of a peer-comparison dashboard are reserved for high-achieving students.

A case study
Students enrolled on the online, part-time MSc in Surgical Sciences at the University of Edinburgh are all trainee surgeons. In response to changes in surgical training that reduced clinical exposure in the workplace, the MSc in Surgical Sciences (Edinburgh Surgical Sciences Qualification, ESSQ) was established in 2007, led by the University of Edinburgh in partnership with the Royal College of Surgeons of Edinburgh. The MSc delivers an
innovative distance learning programme that complements the traditional acquisition of clinical knowledge by surgical trainees, and is the highest recruiting postgraduate course at the University – being taken up by over 1,000 trainees in 70 different countries – and attracts around 100 new students each year (Smith et al., 2013).

The MSc in Surgical Sciences programme utilises a bespoke virtual learning environment (VLE), designed and delivered by the former Learning Technology Section of the University of Edinburgh. Prior to the current study, each assessment metric was displayed on a different page of the VLE, hence there was nowhere for students – or tutors – to see an overall picture of their progress. In accordance with social learning theory (Bandura, 1971), presenting anonymised, comparative progress data to students may increase learner motivation in online programmes, especially true for surgical trainees who are naturally competitive (Hill et al., 2014).

Creating the peer-comparison progress dashboard

During the academic year 2014/15, Tableau® visualisation software was used to create a pilot dashboard which allowed progress data to be displayed on a single page in an effort to provide clear, easy readable and interpretable graphs. Years 1 and 2 of the Masters programme each ran as a course of 20 consecutive modules, with a single, aggregate end-of-year mark. Students were assessed on completion of in-course multiple-choice questions (MCQs), discussion board contributions, essays (Y2 only), and a final written examination. The relative weighting of in-course assessment was 25 per cent in Y1, and 40 per cent in Y2. Included in the dashboard were all the metrics readily captured from the VLE for cumulative modules. These comprised number of logins, number of discussion board posts, percentage score for posts, percentage essay mark, and percentage of the total MCQs completed (a total of 1422 MCQs were released sequentially during the year). Once the template dashboard was created in Tableau®, involving a simple “drag-and-drop” approach to the layout, data were then imported from an Excel spreadsheet file. Subsequent iterations were easy to perform, whereby a single file upload automatically updated the dashboard template. For 2015/16, the pilot dashboard was refined to include a class-ranking tab and a pie-chart illustrating modules included in the analysis, and this version of the dashboard has been used in subsequent years. To view the dashboard, students first had to download the Tableau® Reader desktop application before opening the file sent as an e-mail attachment from their personal tutor to their University of Edinburgh e-mail account. Each student received a unique identification number with the first dashboard in order to view their individual data highlighted from the class data.

Tutor intervention

The data presented in the dashboards were used to provoke e-mail correspondence with students who appeared to be falling behind in their studies. The dashboard was referred to purposely in the opening sentence of personal tutor e-mails to students, e.g. “I am e-mailing you because the student dashboard indicates that you have not contributed very much to the discussion boards to date, nor have you attempted any of the MCQs on the VLE”. In follow-up 1:1 personal tutor meetings, advice was given to these students around equitable participation, addressing potential barriers to discussion board contribution such as feelings of intimidation and/or dominant peers (Symeonides and Childs, 2015). In order to monitor the progress of all students in a cohort and avoid focusing efforts exclusively on poorly performing students, modified e-mails were sent after the release of each dashboard to students who were performing at the required level or above. For example, “Further to the release of the progress dashboard today, I wanted to let you know that so far you are on-target to achieve the 40 per cent threshold required for the in-course assessment. Your average mark this far is $\geq 50$ per cent, so well done and keep up the good work!”. 
Student perceptions
To evaluate the perceived effectiveness of the progress dashboard, a Web-based anonymous survey was issued to students on the programme between 2014/2015 and 2017/2018 (distributed by e-mail). A 3- (yes, unsure, no) or 5-point Likert scale (strongly disagree, disagree, neither disagree nor agree, agree, strongly agree) was used for scoring purposes. The survey consisted of questions relating to the ease of technical use and comprehension of the dashboard, as well as students’ perceptions of its usefulness as a feedback tool and their emotive responses to it (Appendix). The content validity of the survey instrument was evaluated by an external expert (Project Manager of the “Student Analytics VLE Investigation” project 2013–2014, Information Services Group – Technology Enhanced Learning, University of Edinburgh). Survey reliability was assessed using Cronbach’s $\alpha$ to estimate internal consistency between responses to the two questions on “motivation” and “change in study patterns”: $\alpha = 0.75$ indicating satisfactory reliability. In 2017/2018, to incentivise completion of the survey students were invited to enter a prize-draw for a £50 gift-voucher (details entered on a separate page to preserve anonymity). The survey ran for the entire academic year – single completion permitted – following release of the first instance of the dashboard. Student free-text responses were screened to identify common themes until saturation, i.e. no additional themes were found. Institutional ethics approval was not required owing to the anonymous nature of the survey instrument and the voluntary participation of students. Subjective evaluations from students were compared with objective evaluations of full cohort in-course participation. Data relating to discussion board posts and MCQ attempts during 2014/2015–2017/2018 were compared pre- and post-dashboard release using a paired Student’s $t$-test (IBM SPSS Statistics for Windows, Version 24.0). Rank-order change data for low-, middle- and high-achieving students were analysed using one-way ANOVA (GraphPad Prism 8).

Student retention
It was not the express aim of this study to evaluate the impact of the dashboard on students’ retention on the Masters’ programme. However, in order to establish that the dashboard did not have a negative influence on retention, data relating to withdrawal rates across the four years pre- (2010/2011–2013/2014) and post- (2014/2015–2017/2018) dashboard implementation were collated for Years 1 and 2 of the programme (the third year is dedicated to an independent research project), and proportions compared using a $\chi^2$ test (IBM SPSS Statistics for Windows, Version 24.0).

Findings
Dashboards
The progress dashboard was designed to display metrics relating to the continuous assessment components of the Masters programme. A key feature of the dashboard is the interactive element, whereby a student selects their unique, confidential identification number from a drop-down list, and their data then become highlighted from the rest (Figure 1). Average marks for the class are shown as a line in each of the bar graphs. Progress dashboards were released four times in the academic year between November and May. In relation to social learning theory, students have the opportunity to monitor their improvement over time and, with increased course participation, observe better scores and class-ranking (intrinsic reinforcement). Students can click on the data points for top-performing students to reveal activities – number of logins and posts – associated with high scores as an incentive to modify their own study behaviours (observational learning).

Tutor intervention
Progress dashboard-prompted e-mails generated a higher response rate (~40 per cent) from students than previous correspondence (~10 per cent) issued before the introduction of
dashboards. The following e-mail excerpt is a typical example of how students appreciate a “virtual nudge”:

Thank you for your email. Sometimes it’s nice to get a gentle reminder to get jolted out of the rut one finds oneself in.

**Student perceptions**

A total of 137 students responded to the questionnaire, which represents 16 per cent of the total cohorts completing the academic year between 2014/2015 to 2017/2018 (n = 846). The response rate for students in Years 1 and 2 of the programme was comparable: 79/500 (16 per cent) for Year 1, and 58/346 (17 per cent) for Year 2. While the average response rate was 10 per cent between 2014/15 and 2016/17, this rose to 39 per cent in 2017/18 (most likely due to introduction of a gift-voucher prize-draw for completion), and data reveal comparable responses among the four cohorts, diminishing the effect of non-response bias (p-values ranged 0.10–0.26 for first vs last cohort responses, Mann–Whitney U test). It is not clear how many students overall viewed the dashboards since they were issued as an e-mail attachment, i.e. students had to opt-in.

Student feedback provided affirmative endorsement of the dashboard (Table I). The majority (83 per cent) found the data easy to interpret. In total, 94 per cent of respondents found the data relating to their engagement in the MSc very useful (75/137) or somewhat useful (53/137); only 1 out of 137 found them unhelpful (Table I). In total, 55 per cent of respondents said that they will change their approach to study (e.g. completing more MCQs) as a result of seeing the data; 32 per cent do not think that they need to change anything; 13 per cent said they would effectively ignore the data (Table I).

Figure 2 displays the changes in two student engagement measures from first and final release of the dashboard over an academic year. Analysis of which students benefitted from the dashboard, in terms of low-, middle- or high-achievers, was conducted after grouping student data retrospectively based on end-of-year performance quartiles. Although the survey was anonymous, entrants for its associated prize provide insight into the groups viewing the dashboard, i.e. act as a surrogate marker of dashboard use. Of the 50 students who completed
the survey and entered their details into the prize-draw, the majority were in the middle
50 per cent for end-of-year performance scores (Figure 3(a)). The impact of dashboard use on
class rank improvement, based on continuous in-course assessment scores, revealed no
significant difference between low-, middle- and high-achievers ($p = 0.53$, one-way ANOVA)
(Figure 3(b)). These findings are supported by heat map visualisations of student VLE activity;

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes, a lot (n=75)</th>
<th>Yes, somewhat (n=53)</th>
<th>Unsure (n=4)</th>
<th>No, not really (n=4)</th>
<th>No, unhelpful (n=1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Did you find the data relating your course engagement useful?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Did you find the graphs easy to interpret?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Would you prefer the data to be presented in a different format?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Would you like to see additional/different metrics relating to your engagement?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Do you perceive the data provision as an element of feedback?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Will you change your study patterns in light of these data?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. How would you feel about discussion relating to the data within your personal tutor meeting?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table I.** Student responses to three- and five-scale questions

**Notes:** Data are shown as number of respondents, with percentage of total number of respondents in brackets ($n=137$ for Q1 and 3; $n=136$ for Q2, 5 and 6; $n=134$ for Q4 and 7, owing to some skipped questions)

**Figure 2.** Student engagement data

**Notes:** Left: percentage of available multiple choice questions (MCQs) attempted, captured for the first and final dashboard release in an academic year (mean ± SD). Right: number of discussion board posts made each week (mean (range)). Paired Student’s $t$-test ($n=800$, which represents all Y1 and Y2 online students enrolled between 2014/2015 and 2017/2018, excluding those on interruptions of study and withdrawals). *$p<0.01$
top-performing students typically do not reduce their study patterns following dashboard release but those in the bottom 25 per cent are prompted to engage in study (Figure 3(c)).

Student responses were generally positive, as revealed by the free-text quotes given in the questionnaire (Table II). The most common theme to emerge was around peer-performance/class ranking, whereby comparing oneself to others appears to have had a motivational effect, promoting a competitive edge to learners’ engagement. Some students experienced technical issues with the Tableau® software, and several questioned the value of the data displayed.

The top three adjectives selected by students were “interested”, “encouraged”, and “motivated” when asked to describe how they felt when viewing the dashboard data (Figure 4). Despite several students expressing negative emotions, 100 per cent of respondents indicated that they would like to receive regular updates of the student progress dashboard. A plurality (50 per cent) indicated that they would like them made available on a monthly basis, and no-one selected the option “never”.

Notes: (a) Percentage of students within the first, second/third and fourth end-of-year performance quartile who entered the survey prize (n = 50); (b) High-, middle- and low-achieving students’ gain in class rank following dashboard use. Bar = mean (p = 0.53, one-way ANOVA); (c) Example heat map visualisations of VLE activity in the week pre- and post-release of dashboard for Y1 students in the top 25 percent of students (upper panel), and the bottom 25 percent of students (lower panel) for end-of-year performance. Data are displayed as day of week (y-axis) vs hour of day (x-axis); blue dashed-line = the time of completion of survey/access to the dashboard.

Figure 3. The influence of student achievement status on dashboard outcomes
Analysis of free-text comments from the online survey revealed common themes relating to the student progress dashboard

<table>
<thead>
<tr>
<th>Emerging themes</th>
<th>Sample quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Awareness of others in class/relative ranking</td>
<td>“It’s very useful in online education delivery to be able to compare yourself to other students. It has been my worry, that I was not keeping up, or that I was at risk of failure. The data reassured me as to this” “It feels like I’m not alone in this anymore- there are other people and we are all working on the same thing. It’s nice to gauge where I am at compared to others” “Very helpful to see how much work everyone has done as we don’t see each other” “Useful to see each essay scored against other student’s performance. It tells if the essay was generally easy or hard to all of us and where I stand with my performance” “Gives us a chance to understand what other people (aside from just our group) in the course are doing and how you’re tracking compared to the rest of the cohort”</td>
</tr>
<tr>
<td>Motivation</td>
<td>“Getting close to exam period and I have started to prepare for this, I would have increased completion of MCQs regardless of this data but it is useful as a motivational tool to increase engagement” “With it being a distance online learning programme, it’s very easy to get carried away with work and personal life, and unfortunately sometimes end up neglecting ESSQ. I always find that the data which shows me how much I’m engaging either relaxes my mind or pushes me that much more to get through” “It looks as if I’ve fallen behind with contributing to discussion boards. This is something I already knew, but didn’t realise I had fallen behind to that extent! I will make it a point to sit down more to contribute to the discussion board, on top of what I’m already doing” “It was a nice surprise - I wasn’t expecting to be doing so well, so it motivated me to keep it up!”</td>
</tr>
<tr>
<td>Competition</td>
<td>“I think it is easier to engage when there is a little more competition going on” “Knowing our study patterns in relation to other peers are a very interesting mode to generate motivation and competitiveness in the programme we are attaining” “It adds a bit of competitiveness to the whole thing which I think is a good source of motivation”</td>
</tr>
<tr>
<td>Technical issues</td>
<td>“I mean it’s good to get the update. But it’s annoying to install a programme that’s 250 mb just to do this. One which will never be used again” “Large application, needed to clean out some space in my drive for it” “If it could be presented via the university website itself that would be preferable. Large application, needed to clean out some space in my drive for it”</td>
</tr>
<tr>
<td>Data/lag time</td>
<td>“It’s a measure of engagement, not directly related to performance, i.e. a person who’s done more MCQs will score higher regardless if they’ve scored poorly” “I am not sure how accurate the data is as I seem to have done much more MCQs than the graph actually says I’ve done” “Qualitative feedback on my contributions to the board would be far more useful than a graph”</td>
</tr>
</tbody>
</table>

**Table II.**

Student retention
Comparison of withdrawal data pre- and post-implementation of the dashboard reveals a decrease in students leaving the programme post-implementation of the dashboard: 91/755 (12.1 ± 4.9 per cent) vs 75/946 (7.9 ± 2.9 per cent); \( p < 0.05 \). Because this was not a controlled study, it is not possible to attribute the decrease to a causal effect of the dashboard.

Practical implications
The focus of the current study was on student-facing dashboards to allow online students to view their progress in relation to geographically-distant peers, and to provide personal tutors with a visual tool to quickly ascertain an individual student’s engagement with in-course assessment activities. Results herein suggest that providing students with a progress dashboard of cohort comparison data is motivational. While rank-order data can discourage lower performing students (e.g. Wise et al., 2014), the findings of the current study suggest that this did not occur; even the five students who expressed that they felt demotivated by the dashboard data responded that they still wished to receive monthly updates. It is likely that the extent to which a student is motivated to learn is determined by their motivational belief, be it intrinsic or extrinsic (or amotivation). Identifying the motivational characteristics of students is beyond the scope of this study but will be important in strengthening the link between
dashboards and engagement outcomes. A systematic review in 2018 by Jivet and colleagues concluded that peer-comparison dashboards should be used cautiously; contrasting studies found both positive and negative effects of dashboards on student engagement and further research is warranted to understand differences between learners (Jivet et al., 2018). Thus, instructors need to consider carefully the appropriateness of comparison and competition among their particular student cohorts. While such a peer-ranking system may not be appropriate in every setting, all of the students enrolled on the MSc in Surgical Sciences are trainee surgeons and, by definition, they have entered a very competitive field and are familiar with class-ranking from their undergraduate medical education.

The positive response from students on the perceived usefulness of dashboards is in agreement with the literature (Verbert et al., 2014; Yoo et al., 2015). For a better measure of their effectiveness, Bodily and Verbert (2017b) recommend asking students about the perceived effect on behaviour/study patterns. In this study, the majority of respondents indicated that they will change their approach to study in light of the dashboard data. Use of progress dashboards can stimulate participation in online assessment activities, as this student quote from the current study reveals:

I now have a greater appreciation of assessment weighting and will prioritise case discussion posts over completing the module.

Owing to the anonymous nature of the study, it is not possible to verify whether students did increase their in-course participation as implied. However, an objective analysis of MCQ completion and discussion board participation following release of the dashboard supports a positive change in students’ study behaviours. Furthermore, subgroup analysis of survey prize entrant data to ascertain the impact of the dashboard on academic performance of individual students reveals no difference between low-, middle- and high-achievers in terms of class-rank gain.

Some students expressed doubts about the accuracy and value of the data. This can be explained by the lag time in data capture and dashboard release in relation to the current module running in the programme. As a response to confusion expressed by a few students around which data account for their progress view, the pie chart depicting specific modules included in the dashboard was added after the first year of piloting the dashboard.
Park and Jo (2015) stress the preference for dashboard data to be objective and trustable, and not related to any kind of evaluation. It was made clear to students on the MSc in Surgical Sciences that the dashboard was only released to students and the personal tutor and did not form part of any evaluative process.

**Tutor intervention**

This case study used the dashboard as a fast tool for identifying those students falling behind in their studies, as measured by their level of VLE engagement and marks for in-course assessment, and then intervening via e-mail to alert the student of their position, and discuss the reasons for their performance, and advise on ways to improve. Interestingly, progress dashboard-prompted e-mails have generated a higher response rate from students than previous correspondence simply commenting on their progress, and may be viewed as a proxy to changing behaviours. This possibly relates to the visual nature of a graph carrying more authority than if the data were presented as text, a list or table. Indeed, it has been shown that people are more likely to be persuaded by graphical representations of data compared to the same information in text format (e.g. Nyhan and Reifler, 2018). In the learning analytics arena, communicating progress data to students as dashboards of charts allows for ease of transmission, but there is a risk that graphs could be misinterpreted by some and carry disproportionate meaning (Alltree et al., 2014). With that awareness, tutors must ensure supportive lines of communication exist between themselves and the student to avoid a negative graphic of progress resulting in disengagement with the learning environment.

The dashboard has acted as a conversation catalyst during tutor-student meetings. For a student to see where they rank in relation to their peers, and the class average, makes for a more meaningful discussion on their progress and setting achievable goals. Prior to the dashboard, tutors would have to check on multiple pages within the VLE and on spreadsheets offline in order to build up a picture of how a student was engaging with the programme. Now, the dashboard provides an instant view, which is especially convenient if a student chooses to telephone their tutor on-the-fly to discuss their studies. This is important since accessibility of the personal tutor is a key determinant of student uptake (Walsh et al., 2009).

**Student retention**

While there was an appreciable “drop-out” rate from the programme each year (5–15 per cent), this is less than that observed in other distance learning masters programmes; (e.g. 40 per cent reported by Patterson and McFadden, 2009). The higher retention rates for the online MSc in Surgical Sciences likely reflect the programme attracting competitive, highly motivated and dedicated surgical trainees who demonstrate study behaviours that promote successful academic outcomes (Stienen et al., 2018). The introduction of a progress dashboard was associated with an average 4 per cent decrease in number of student withdrawals on the MSc in Surgical Sciences. This is in agreement with findings from other institutions (e.g. Arnold and Pistilli, 2012), although the confounding effect of additional factors unrelated to the dashboard cannot be ruled out. Since this was an observational study, not involving a control/intervention group comparison, it would be naïve to attribute the improved retention solely to the dashboard/tutor intervention. Nevertheless, it is clear that withdrawal figures did not rise with the introduction of progress data, thus countering concerns around their potential negative impact.

**Study limitations**

The low response rate to the questionnaire may have resulted in non-response bias; there is a potential for high performers to view the dashboard attachment and complete the survey. However, entries received for the voucher prize in 2018, and dashboard-prompted e-mail correspondence from students to their personal tutor in the earlier years, suggest students from a wide range of rankings participated in this study.
Dashboards were released to students via e-mail attachment. A major uncertainty in this study is the relative use of University e-mail accounts by students. A recent literature review of student-facing learning analytics reporting systems reports a low use of dashboards by students (Bodily and Verbert, 2017a), and its authors advocate future research to examine how to increase student use of such tools. Acknowledging student feedback on technical issues downloading the necessary software, steps to incorporate the dashboard tool into the VLE in future iterations are justified, but the primary reason for e-mail release was to ensure supportive communications accompanied the dashboard.

The survey data collected in this study represent static data, i.e. responses at a single time point were captured. Dynamic relationships between the dashboard data and student views would be an interesting aspect to pursue future iterations, since students are likely to have varying perceptions of their progress and motivation throughout the course (Pardo et al., 2017).

Conclusion

Presenting anonymised comparative progress data to students may increase learner motivation in online, distance learning programmes, especially true for surgical trainees who are naturally competitive. Students value being able to see their ranking alongside geographically disparate peers and it can incentivise a change in study behaviours. Progress dashboards provide a convenient and efficient means for personal tutors to monitor students’ progress quickly, at a glance, and can be used as a catalyst to identify “at risk” students and affect interventions in order to assist students in taking greater ownership of their learning and facilitate achievement of their potential. Clearly, there are differences in learner motivations resulting from dashboards and a subset of students for whom this is particularly useful. Future research is recommended to identify the characteristics of students who will benefit most from peer-comparison dashboards and to design effective alternative strategies for others to best monitor their progress.

References


References


## Appendix 1

<table>
<thead>
<tr>
<th>Student feedback LA project Y1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Student progress dashboard – ESSQ</td>
</tr>
</tbody>
</table>

Thank you for taking the time to complete this short survey.

Your opinions and feedback are extremely important in shaping delivery of student ‘progress dashboards’ on the ESSQ programme. Your responses will be anonymised.

1. Did you experience any difficulties downloading/using the Reader software?

   - [ ] Yes
   - [ ] No

   **Comments**

2. Did you find the data relating your eeSURG engagement useful?

   - [ ] Yes, a lot
   - [ ] Yes, somewhat
   - [ ] Unsure
   - [ ] No, not really
   - [ ] No, unhelpful

   **Please explain why**

3. Did you find the graphs easy to interpret?

   - [ ] Yes
   - [ ] No
   - [ ] Unsure

   **Please explain why**
4. Would you prefer the data to be presented in a different format?
   - Yes
   - No
   - Unsure

   If yes, please describe your preferences:

5. Would you like to see additional / different metrics relating to your ESSQ engagement?
   - No
   - Yes

   If yes, please say what:

6. How did the data make you feel? Tick as many as apply.
   - Surprised
   - Worried
   - Confident
   - Interested
   - Anxious
   - Proud
   - Encouraged
   - Relieved
   - Concerned
   - Re-assured
   - Motivated
   - Demotivated

   Other (please specify):
7. Will you change your study patterns in light of these data?
   - Yes, I will most likely change
   - No change...effectively ignore
   - No I don't think I need to change

   Please describe your new approach:

   [Blank space for description]

8. How frequently would you like to receive such data?
   - Never
   - Monthly
   - Weekly
   - Daily
   - After each module ends
   - On-demand

9. Do you perceive the data provision as an element of feedback?
   - Yes
   - No
   - Unsure

   Any comments?

   [Blank space for comments]

10. How would you feel about discussion relating to the data within your personal tutor meetings?

<table>
<thead>
<tr>
<th>No problem!</th>
<th>Neutral</th>
<th>Unsure</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
</tbody>
</table>

   My reaction is:

   Any comments?

   [Blank space for comments]
11. Would you like to discuss your progress with your personal tutor now?

☐ Yes - please e-mail paula.smith@ed.ac.uk

☐ No

12. Currently, this dashboard is only available to students (anonymised) and Dr Paula Smith, Academic eFacilitator. Who else do you think should be given access to it? Tick as many as apply.

☐ Personal Tutor

☐ Year Director

☐ Programme Director

☐ University staff, e.g. Academic Services

☐ No-one else

Other (please specify)

☐

13. If you have any further general comments or suggestions to help improve how we present progress data, please leave them here.

☐

Thank you for completing this survey. Feel free to e-mail paula.smith@ed.ac.uk if you have any queries.

If you would like to enter a prize draw to win a £50 Amazon voucher, please click on the Next button below and you will be taken to a separate screen to leave your details.
### Student feedback LA project Y1

2. Chance to win a £50 Amazon voucher

14. Please leave your name and/or matriculation number to be eligible to enter the prize draw.

```

```

---

**Corresponding author**
Paula Smith can be contacted at: paula.smith@ed.ac.uk

---

For instructions on how to order reprints of this article, please visit our website: [www.emeraldgrouppublishing.com/licensing/reprints.htm](http://www.emeraldgrouppublishing.com/licensing/reprints.htm)
Or contact us for further details: permissions@emeraldinsight.com
Framework to support personalized learning in complex systems

Aklilu Tilahun Tadesse and Pål Ingebrigt Davidsen
Department of Geography, System Dynamics Group, University of Bergen, Bergen, Norway

Abstract

Purpose – Numerous studies document that students struggle to comprehend complex dynamic systems (CDS). The purpose of this paper is to describe a design framework applied to the creation of a personalized and adaptive online interactive learning environment (OILE) to support students in their study of CDS.

Design/methodology/approach – A holistic instructional design is applied in five steps to create the OILE. The OILE has the following characteristics: first, it presents a complex, dynamic problem that learners should address in its entirety. It then allows learners to progress through a sequence of learning tasks from easy to complex. Second, after completion of each learning task, the OILE provides learners with supportive information based on their individual performance. The support fades away as learners gain expertise. Third, the OILE tracks and collects information on learners’ progress and generates learning analytics. The OILE was tested on 57 system dynamics students.

Findings – This paper provides evidence that supports the theoretical design framework from the literature. It also provides a sample from students’ progress logs to demonstrate how the OILE practically facilitated students’ cognitive development. In addition, it provides empirical evidence regarding students’ attitudes toward the OILE that was obtained from administering two questionnaires.

Originality/value – In light of supportive evidence from the literature, students’ progress in the cognitive domain, and confirmative response in the affective domain, the use of personalized and adaptive OILE to support learning about CDS is considered promising.

Keywords System dynamics, Learning environment, Systems thinking, Adaptive learning, Learning analytics, Personalized learning

Paper type Research paper

1. Introduction

Decision makers and people in general face a wide range of increasingly complex, dynamic problems in both the public and private sectors (Sterman, 1994; Davidsen, 1996; Jonassen, 1997; Barlas, 2007; Greiff et al., 2012). These problems have a dynamic nature (change over time) and they commonly originate from the internal structure of the system that generates the problem (Diehl and Sterman, 1995; Davidsen, 1996). Structure is the cause and effect relationship among variables that define the system. Numerous studies document that people have difficulties comprehending complex dynamic systems (CDS) and managing these systems effectively (Dörner, 1996; Moxnes, 1998, 2004; Cronin et al., 2009).

Moxnes and Saysel (2009), in their research about misperceptions of global climate change, show that people have cognitive difficulties with building appropriate mental models unless they are supported by proper strategies. Other studies on renewable resource management (Moxnes, 1998, 2004; Jensen, 2005) and on global warming (Sterman and Sweeney, 2007) have confirmed the problem.

The difficulty in understanding CDS arises from the lack of three different types of capabilities: cognitive capability to comprehend structural complexity; skills to infer the dynamic behavior of a system from its underlying structures; and effectiveness of methods, techniques and tools that are available to analyze systems (Davidsen, 1996; Spector and Anderson, 2000; Jonassen, 2000; Ifenthaler and Eseryel, 2013; Van Merriënboer and Kirschner, 2017).
The objective of this paper is to develop a framework that identifies educational methods that consist of a well-composed set of instructional techniques and tools. The techniques are manifested in the form of educational tools. The tools are those created to materialize the instructional techniques that support students in their study of CDS.

The paper describes a design framework applied to the creation of a personalized and adaptive online interactive learning environment (OILE) to support students’ learning. Often the terms “personalized learning” and “adaptive learning” are used in similar educational contexts, referring to “the tailoring of education to learners’ current situation, characteristics and needs” (Graf and Kinshuk, 2012, p. 2592). However, each term has its own specific definition and focus area. The US Department of Education (2017) defines personalized learning as “an instruction in which the pace of learning and instructional approach are optimized for the needs of each learner” (p. 9). Graf and Kinshuk (2012) share a similar essence in their definition of personalized learning – “tailoring education to learners’ current situation, characteristics, and needs in order to help them to achieve the best possible learning progress and outcomes” (p. 2592). Here, the focus is more on “the consideration of the learner as an individual person.” On the other hand, the focus in adaptive learning is on “the aspect of achieving the tailoring automatically (typically by a learning system).” Bilic (2015) defined adaptive learning as the “adjustment of one or more of the characteristics of a learning environment” to match the needs of learners. The adjustments can be on how the learning tasks are presented, on how they are sequenced depending on the learners’ progress, and on how supportive information is provided. While adaptive learning can also be applied to groups of learners – adjusting the learning environment to those groups, personalized learning always “focuses on the individuals, regardless of the fact whether they work alone or in groups” (Graf and Kinshuk, 2012, p. 2592).

Personalized and adaptive OILEs are important to students of educational programs that employ system-thinking concepts. Courses offered in such educational programs are interdisciplinary and often breakdown the barriers among fields such as natural science, engineering, political science, economics, and medicine. Furthermore, students who register for such programs often come with diverse academic and cultural backgrounds and experiential perspectives. Such issues create practical difficulties for educators when trying to bring the students to the same level and help them proceed at an equal pace. This is because:

1. Individual students are unique and have unique learning paths.

2. Educators in schools and universities seldom educate their students to be “problem solvers to solve problems that emerge from their disciplines.” Rather they teach them about the disciplines. Students are taught about “sociology, psychology, history, biology” not “how to be a sociologist, psychologist, historian, or biologist” (Jonassen, 2010, p. xxi).

3. Some students are very shy to come forward and ask for help either from their teachers or from their colleagues due to cultural and experiential perspectives. It is difficult for a teacher to identify students who are struggling with their learning material unless either the students go to the teacher on their own or exam results are published.

However, personalized and adaptive OILEs can support students while studying CDS in their own time and their own pace. In addition, they can help to minimize gaps that exist among students by providing computer-mediated resources based on the unique needs of learners. Furthermore, they provide predictive learning analytics to the teachers about the progress of each student. They also help teachers and instructional designers to identify
learning materials and learning tasks that need revision to improve the quality of the teaching/learning.

2. Research questions

This paper aims to demonstrate how the OILE could be designed to support individual students in their study of CDS. In particular, the following three research questions were investigated:

RQ1. How can one address the cognitive challenges associated with the teaching/learning of CDS?

RQ2. Can personalized and adaptive OILE facilitate students’ cognitive development in their study of CDS?

RQ3. Can personalized and adaptive OILE support students’ affective domains of learning in their study of CDS?

3. Characteristics of the OILE

This section of the paper describes the characteristics of the OILE created to address the challenges associated with the teaching/learning of CDS. The OILE has the following three predominant characteristics:

1. It presents an authentic, complex, dynamic problem that the learner should address in its entirety. It then proceeds to allow learners progress through a sequence of learning tasks from easy to complex.

2. When solving the problem, learners interact with the OILE. After completion of each learning task, the OILE provides learners with supportive information based on their individual performance. The support fades away as learners gain expertise.

3. It tracks and collects information on students’ progress and generates learning analytics, which are used to assess the students’ learning.

The figures show sample screenshots associated with the characteristics offered by the OILE. Figure 1 is a sample screenshot from the OILE’s welcome page and Figure 2 is a sample screenshot from item description page. The arrows in Figure 2 indicate the components of an item description page (detailed explanation is given in section 4.5 – tools and interface).

Figure 3 represents a learning path of a student who has worked on the OILE from question number “Q1-1” to “Q2-10”. The green lines represent progress in performance and the red lines represent movement to remedial questions. Detailed discussions on learning paths in the context of this paper are given in Sections 4.4 and 5.1.

Designing interactive learning environments that effectively support learning in and about CDS is a challenging but fascinating task, which requires the synthesis of different instructional methods, techniques and tools (Sterman, 1994; Davidsen, 2000; Eseryel et al., 2011). Because these learning environments are required to influence the formation of mental models that govern learners’ decision making and action in CDS (Davidsen, 2000). Moreover, the learning environments are required to provide contexts to practice scientific methods in both virtual and real worlds, while facilitating the practice (Sterman, 1994). A review of the literature in the area of supporting learning in and about CDS show but a few interactive learning environments that have been designed to offer such a comprehensive support.

Most of the interactive learning environments designed to foster learning in and about CDS are ready made black-box simulators with user-friendly interfaces that
display the surface relationship between input provided by the user and output provided by the simulation engine (Alessi, 2000; Pavlov et al., 2015). Such learning environments provide little or no information about the internal structure of the CDS. They provide platforms either to conduct controlled experiments (trials and occasional success) as in the case of management flight simulators or to train learners with specific tasks/procedures, for example, driving a car or flying an aircraft without diving into the inner workings of the devices (Sterman, 1994; Alessi, 2000; Pavlov et al., 2015). Moreover, these learning environments do not require learners to pass through rigorous scientific methods such as problem identification, hypothesis formulation, analyses and interpretation of results. On the other hand, there are few other interactive learning environments that have been designed to promote discovery learning, where learners are
engaged in a scientific inquiry to uncover the underlying structure of CDS and solve their associated problems (Alessi, 2000). However, what is often missing in these kind of learning platforms is an effective learning support that facilitate the students’ learning. Even if there are support mechanism in these platforms, often this support has been designed to fit the so-called average learners without accounting for strugglers and/or top performers.

That said, there are system dynamics based interactive learning platforms designed and implemented in Forio Simulate (forio.com) for courses offered at Sloan School of Management (see Sterman, 2014a, b) and at a number of other major US universities. These learning environments have been designed to offer support to learners that study key concepts in strategic management, system dynamics and related fields. They use real world case studies and allow students to carry out experiments using simulation models so as to gain experience with complex systems. With the help of their user-friendly interfaces, learners have the capability to alter the various settings of the simulation, make decision using input devices and explore the consequences of their decisions. However, these platforms do not offer support for students to (re)create the simulation models that represent the structure of the underlying CDS. They also do not provide thorough and comprehensive support to individual students, at each stage of the learning process, to help them uncover the relationship between the underlying structure of each CDS and its development over time (i.e. the relationship between structure and behavior). Often debriefings are provided merely toward the end of the learning process.

The design framework proposed in this paper, however, aims at incorporating these elements of learning that are missing from the learning process. This framework allows learners to be engaged in a scientific discovery practice to address authentic learning tasks, uncovering the relationship between structure and behavior of CDS. The OILE

![Sample learning path of a student](image-url)

**Notes:** The green lines represent progress in performance and the red lines represent movement to remedial questions; the yellow arrow indicates the tasks sequence.
provides intermittent scaffolding to support learners based on their individual performance so as to inform the development of both their mental models and the formal models they would create. The scaffold fades away as learners gain expertise. It also collects learning analytics using built-in trackers to inform both teachers and students regarding the learning progress and allow for the redesign of the instruction when needed.

The section below presents the framework used to design the OILE as well as the underlying instructional design theories, where the three characteristics of the OILE are distilled out. It then presents the instructional method as well as the instructional techniques employed in the design of the OILE to support students’ learning. Finally, the section presents the organization of the learning tasks in the OILE.

4. A five-step design framework for the OILE

The OILE is designed under the guidance of a five-step design framework listed below:

1. identify instructional design models;
2. identify authentic (real world) learning tasks;
3. identify instructional methods;
4. identify instructional techniques; and
5. design interface and implement the tool.

Table I summarizes the five steps involved in the design of the OILE, and the subsections below give a detailed account of the five steps. The analyses presented under these five steps give a literature based theoretical foundation to the OILE development and answer the first two questions of the research:

1. How can one address the cognitive challenges associated with the teaching/learning of CDS?
2. Can personalized and adaptive OILE facilitate students’ cognitive development in their study of CDS?

4.1 Identify instructional design models for the teaching/learning of CDS

The systems thinking approach is, the whole is always more than the sum of its parts, indicating that the structure of parts synergizes to produce the resulting dynamics of a system.
Hence, the instructional design models considered in the design of the OILE should foster this holistic perspective.

Van Merriënboer and Kirschner (2017) argue that there are three common problems that instructional designers should address when designing instructional models that support learning in and about complex systems, “compartmentalization of knowledge, skills and attitudes; fragmentation of what is learned in small, incomplete or isolated parts; and the transfer paradox” (p. 13). They, however, claim instructional design models that follow the holistic perspective can solve these problems by integrating the different domains of learning into a unit of instruction and by facilitating the development of an integrated knowledge base that increases the chance of transfer. Spector (2000) also noted the above-mentioned problems and proposed a set of five basic principles that complements the holistic instructional perspective, to consider during instructional design (p. 524):

- Learning principle (L) – learning is fundamentally about change;
- Experience principle (E) – experience is the starting point for understanding;
- Context principle (C) – context determines meaning;
- Integration principle (I) – relevant contexts are broad and multi-faceted; and
- Uncertainty principle (U) – we know less than we are inclined to believe.

This study utilized these five basic principles and the holistic perspective as inclusion criteria for choosing instructional design models that informed the creation of the OILE.

The first column of Table I shows the six instructional design models that met the inclusion criteria and that have influenced the development of the OILE: the 4 C/ID, first principles of instruction, CLE, TCI, cognitive apprenticeship and Elaboration theory.

The primary emphasis in learning environment design varies across the six models. Nevertheless, the models have four key features that make them suitable for the design of learning environments intended to foster understandings of CDS.

First, they offer a unifying perspective regarding learning tasks. These instructional design models argue that the learning tasks should:

- be at the center of the instructional design;
- be based on authentic problems;
- comprise the entire knowledge and skills that learners would be able to acquire when they complete the entire learning tasks;
- be designed in a way that learners can address the authentic problem in its entirety, from “start to finish, rather than discrete pieces” of the problem; and
- be designed in a way that learners can progress from simple to complex steps in their analysis of the entire task.

These core principles constitute the foundation of the first characteristic of the OILE (Section 3) – the OILE first presents an authentic, complex dynamic problem then allows learners to progress from simple to complex steps in their analysis of the entire problem.

Second, the instructional design models underscore the importance of providing instructional scaffolding that gradually fades away over time as learners gain expertise. The authors of the six instructional design models argue that learners should receive the right support at the right time. The support bridges possible learning gaps and sustains students’ engagement with their learning tasks. This principle serves in part as a base for the second characteristic of the OILE – designing supportive information that fades away,
and partly as a base for the third characteristic of the OILE – the collection of process log information to evaluate the status of the learners.

Third, they all promote holistic instructional design. They recognize the dynamic interdependency between the elements that constitute an instructional system of complex learning that makes the instructional system an irreducible whole (Van Merriënboer and Kirschner, 2017). For example, an analysis of the performance of a learner regarding a certain learning task is used to design the supportive information presented to the learner. This information is then used as an input to design the next task. However, the level of difficulty of the next task is determined based on the learner’s performance while addressing the previous task. The learner can subsequently progress either to more complex or to a simpler learning task. This key principle served as a foundation to all of the three characteristics of the OILE.

Fourth, they emphasize the importance of transfer of knowledge and skills to everyday life. The tasks the learners undertake as part of the learning experience and the instruction they follow in the learning environment should help the learners transfer their knowledge and skills to related real world settings. This last principle is also integrated into all of the three characteristics of the OILE.

4.2 Identify authentic learning tasks

The above-mentioned six instructional design models require authentic learning tasks to be presented to learners, and so does the design framework of the OILE.

An authentic, complex and dynamic problem has been identified during the design of the OILE. The learning task is a case study to be completed by master students in the system dynamics program at the University of Bergen, Norway. The case study is about Mr Wang’s Bicycle Repair shop, aimed at teaching the students about the causes of oscillation.

Oscillation is one of the fundamental modes of behavior produced by non-linear feedback systems. It occurs in virtually all business areas such as in commodity markets, labor supply chains, manufacturing supply chains, and the real estate market. Using this case study, the students investigate the causes of oscillation in the backlog of the Bicycle Repair shop. In this study, backlog represents the bicycles that have been delivered to the company and that are in need of repair.

The content of the learning environment has been reorganized based on the recommendation of “best practices in modeling” by Martinez-Moyano and Richardson’s (2013) and Richardson’s (2014a, b, c) “canonical sequence” framework that help to lead learners seamlessly from problem identification and conceptualization to model testing and evaluation. The objectives at each stage of the learning and the intended learning outcomes were formulated with the help of frameworks provided by Munoz and Pepper (2016) and Schaffernicht and Groesser (2016).

The Mr Wang case study is divided into five tasks. A task is a subset of challenges with specific objectives that students should be able to achieve upon completion of that task. The first task of the case study focuses on problem identification and definition. The learners are required to identify the problem the repair shop has. Tasks 2–5 concentrate on hypothesis formulation and analysis of that hypothesis. The students are required to formulate a hypothesis about the underlying causal structure of the problem. They then proceed to analyze the relationship between that structure and the consequent dynamic behavior by building computer models. The students carry out this task in a reiterative process until they arrive at a structure that best explains the identified problem. The complexity of the underlying causal structure and its analysis increases as student progress from Tasks 2–5.
4.3 Identify an instruction method

Scaffolding is the main instructional method applied in the design of the OILE. Wood et al. (1976) defined scaffolding as a “process that enables a child or novice to solve a problem, carry out a task or achieve a goal which would be beyond his unassisted efforts” (p. 90). The support provided is “meant to extend students’ current abilities” so that they can carry out the “bulk of the work required to solve the problem” (Belland, 2017, p. 17).

The scaffolding instructional method comprises three elements: dynamic assessment, provision of just the right amount of support and intersubjectivity (Belland, 2017). Dynamic assessment and provision of just the right support are interrelated and must be carried out iteratively (Belland, 2017). The dynamic assessment determines whether the learners are constructing knowledge and skills from the learning tasks and whether they are on the right path to be able to perform the tasks independently. If the assessment indicates that the learners are struggling to make meaningful learning, the scaffold level increases. If the learners are on the right path, the scaffold gradually fades away (Wood et al., 1976; Belland, 2017).

Intersubjectivity refers to a shared understanding between the “scaffolder” (teacher or learning environment) and the “scaffoldee” (learner) regarding a successful performance of a learning task (Belland, 2017). It is very important that, when learners work independently at the end of the instruction, they should be able to recognize whether they are doing so correctly or not (Wood et al., 1976; Wertsch and Kazak, 2005; Belland, 2017). Intersubjectivity is crucial in building the learners’ self-efficacy, i.e. “the context-specific belief that one can perform successfully” (Myyry and Joutsenvirta, 2015, p. 121). The third and fourth columns of Table I present these three features of the scaffolding instructional method and the instructional techniques applied to manifest the three features.

Instructional scaffolding is a perfect candidate for “casting” the complex and dynamic problem via the OILE. This is because it:

- comprises most of the key features of the six instructional design models such as task (content) centered design, simple to complex progression of learning tasks with a holistic perspective, provision of the right support when needed, and focus on transfer of knowledge and skills into real world tasks; and
- requires the learners to complete the entire learning tasks, which consist of all the knowledge and skills the learners need to gain to become more expert.

4.4 Identify instructional techniques

4.4.1 Instructional techniques for dynamic assessment. Dynamic assessment is done using multiple-choice questions (MCQ) and open-ended questions (OEQ). The format of the MCQ is consistent throughout the learning environment with four alternatives, except in two specific questions that have only two alternatives. There is only one correct choice per question and learners can only choose one answer at a time. In the OEQ format, learners are asked to predict behavior over time graphs. The students draw their predictions on the OILE and submit their answers.

The learning material is designed in a way that, at each stage of the learning activity, learners are posed a question to solve a problem. The learners work on the question and give their response either by choosing one of the MCQ alternatives or by drawing behavior graphs. The questions range from identifying a vivid problem to hypothesizing a causal structure responsible for that problem and analyzing the relationship between the suggested structure and the consequent dynamic behavior by building computer models. During the model-building phase, learners are required to work with a modeling software installed on their local computers. In this stage, the learners are often required to switch between the OILE and the modeling software, so that they can have hands on activities.
The questions are arranged in sequences called “learning paths.” A learning path is a sequence of questions that learners pass through, while working on the complex and dynamic problem on their own pace and time (see Figure 3). Each learner has her/his own unique learning path. In general, there are linear and branching sequence questions in the learning path of a learner. Linear sequence questions are those where a learner moves to the next question after finishing the previous question without any precondition. Branching sequence questions are those where the next question depends on the performance of the learner in the previous question. The branching technique is discussed in detail in the next subsection.

4.4.2 Instructional techniques for provision of just the right support. Five instructional techniques have been used in the OILE to “provide just the right support”: storytelling, repeated trial, feedback, feed-forward and item branching.

The storytelling technique is used to present the content of the learning material. It is used to contextualize the students’ learning. This technique helps to link what students already know with the new information. It also helps to provide important information to learners that help them solve the learning tasks. The storytelling technique is used in almost all of the six instructional design models under different names: adjusting scaffolding (Jonassen, 1999; Reigeluth et al., 2017), supportive information (Van Merriënboer and Kirschner, 2017) and activation of prior knowledge (Merrill, 2013; Francom and Gardner, 2014).

The repeated trial technique is used to give students the opportunity to try a question multiple times. The students have three chances to try to answer a MCQ that has four alternatives. This technique helps to design different levels of support for the students. A student who has failed to respond correctly to a question twice receives more support than those who have failed only once.

Repeated wrong choices of students’ serve as good indicators for possible misconceptions, which the teachers can address during face-to-face instructions. Also, they help the teachers and the instructional designers to identify learning tasks that need revision to improve the quality of the OILE. From the students’ perspective, the repeated wrong answers help them recognize their performance level and their progress in the learning tasks.

The third and fourth instructional techniques are provision of feedback and feed-forward. Gagné (1985) underscores that provision of timely and informative feedback is crucial for learning. In the OILE, every time students respond to a question a feedback page opens. If the question is an OEQ, then the students receive suggested answers so that they can compare their response with the suggested answers.

However, in the case of MCQ, the students receive different feedback based on their individual performance. If the chosen alternative is correct, the students get the reason why that alternative is correct and why the other alternatives are wrong. Such feedback has two objectives: to make sure that the students know the correct reason that their responses are correct, thereby strengthening intersubjectivity between the scaffoldee and the scaffoldor; to prevent the impact of “guessing” on subsequent tasks, thereby serving as a feed-forward.

If the students’ responses are wrong and are not their third trials, they get either a corrective feedback or an ordinary feedback with item branching.

A corrective feedback is a feedback that explicitly shows the reason why the students’ answers are wrong. Whereas an ordinary feedback with item branching is a feedback that simply tells the students their reply is incorrect. Unlike the corrective feedback, the students do not receive information about why their answers are wrong. Rather, the students are asked to branch to tasks that are easier than the previous but under the same conceptual framework so that they can figure out on their own why their previous responses were wrong.
The option for providing either “corrective feedback” or “ordinary feedback” with item branching depends on individual students’ learning paths and the stages at which they are in the learning tasks.

Students receive corrective feedback while they are in the early stages of problem identifications. Students’ continued engagement in the learning environment can be affected by their early perception regarding what they are going to do (Jonassen, 1999). The learning tasks and the support provided during the early stages of the learning process should help the learners understand the problem statements clearly.

Corrective feedbacks gradually fade away and are replaced by ordinary feedbacks with item branching as students advance through the learning material. Students are allowed to branch up to three levels. However, if the students fail to respond correctly at the lowest level, they will get corrective feedback. If the students are successful in responding to the lower level questions, they will move up to higher levels and work again on the questions they failed to respond correctly. In doing so, the students move up and down the ladder. If the students respond correctly to the questions at the top level, they progress to a relatively complex task. As the learners progress through the learning material and gain more expertise, the item branching reduces from three levels to two levels and finally to one.

4.4.3 Instructional technique for intersubjectivity. In the OILE, intersubjectivity is maintained by allowing students to pass through iterative steps at each stage of the learning material and by providing summaries at the end of certain group of learning activities.

Every time students are asked to identify a problem, for example, they will be asked first to identify the variable that represents the symptom of the problem (the stock/accumulator). Then they will be asked to identify the variables that cause the stock/accumulator to change (flows) and finally variables that influence the flow rates to change (auxiliary variables or parameters). Van Merriënboer and Kirschner (2017) classify such skills as “recurrent constitute skills.” These “recurrent” skills could be acquired either by following certain procedures and/or rules or by “continually practicing them in order to automate those constitute skills” (p.97). However, skills such as identifying a stock or a flow variable from a problem description are achieved by building schema of those variables (Jonassen, 2000). Van Merriënboer and Kirschner (2017) classified these skills as “non-recurrent constitute skills.” The techniques used in the OILE help to strengthen the construction of both “recurrent” and “non-recurrent” constitute skills, thereby establishing intersubjectivity between the scaffoldee and the scaffoldor.

During the model behavior analysis stage, students are asked to chop time on the basis of monotonic behavior developments, so that they can explain each monotonic development referring back to the structure of the model. In doing so, students practice and strengthen their analytical skills. Van Merriënboer and Kirschner (2017) call this “part–task practice.” The main objective of the part-task practice is to develop particular sub-skills to acquire automatic performance (p. 49).

To reinforce certain subtle concepts of the learning material, summaries are provided at the end of related learning activities. In the summaries, the main insight of each learning task is highlighted so that students can easily integrate the new concepts with the learning objectives.

4.5 Design the interface and implement the tool
The OILE is a web-based instructional tool designed to manifest the instructional method and techniques described in the previous subsections. The OILE is designed around the Mr Wang case study.

The five tasks of the case study (see section 4.2) are organized under three different OILEs for ease of management. The first OILE consists of Tasks 1 and 2. The second
consists of Tasks 3 and 4 and the third OILE consists of Task 5. However, all three OILEs have the same characteristics and they have similar interface design.

Under each OILE, a task is further divided into items. An item is a specific challenge presented to a student. An item comprises stimulus material and a question (see Figure 2). The stimulus material manifests the storytelling technique (section 4.4) and is a description of a problem a student is supposed to solve. Stimulus materials are presented as text, behavior graphs, causal loop diagrams, and/or built-in models. Simulation buttons and time series graphs that display computer simulation results accompany built-in models. The stimulus material gives context for a challenge presented in the form of a question. A question is a specific challenge a student is supposed to solve on the basis of the stimulus material. The case study consists of 105 questions in total.

Two user interface design formats have influenced the interfaces of the OILEs: the user interface design of OECD (2013) and the storyboard format of Alessi and Trollip (2001). The OILE’s interface is divided into different pages: a welcome page, task description page, item description page and supportive information provision page.

Figure 1 shows a screenshot of a sample welcome page. The welcome pages introduce the general objective of the case study. Navigation buttons placed at the top and bottom right of the screen guide the students’ movement through different pages. Buttons that lead to the immediate next page (default next page) are highlighted with bright green colors. However, students have the option to go to pages other than the “default next page” using navigation buttons that are not highlighted. For example, a student who comes from a task description page to the welcome page for the second or third time can jump directly back to the task description page (where the student was) by clicking on the appropriate navigation button (last visited page).

Figure 2 shows a screenshot of a sample item description page. In the item description page, stimulus material appears in the top part of the screen. The question appears in the lower part of the screen, and borders visually separate it from the stimulus. The top left of the screen displays the name of the case study.

The navigation buttons placed at the top right of the screen guide the students’ movement through different pages. Item and task numbers are shown at the top left edge of the screen. The multiple-choice alternatives are placed at the bottom left edge of the screen.

The OILE is built on an interface of a computer modeling software, Stella Architect version 1.4. The OILE has been hosted on the isee exchange online platform (https://exchange.iseesystems.com/login). It can run on any web browser. However, students need special permission to access the OILE page.

Once students log into the OILE page, a special tracker built on the Stella Architect software tracks their process log information. The tracker records information such as name, students’ performance in the case study, the pages the students have navigated and the amount of time they have spent on a page. It also records what kind of support the students have received. The process log information was collected in the form of comma separated values (csv) files and time series graphs. Sample time series graphs and csv files are presented in Figure A1 and Table A1, respectively.

The tasks in the OILE require students to engage in hands on modeling exercises. Hence, students are required to install modeling software such as Stella Architect into their local computers so that they can easily switch between the online material and their modeling software.

5. Sample results and discussion
October 2017 and 2018, first year system dynamics master program students (57) at the University of Bergen, Norway, used the OILE to carry out the Mr. Wang Bicycle Repair Shop case study, a part of the students’ mandatory course work.
In its original pencil and paper format, students were introduced to the case study by a professor. The students submitted their work after working for a week. In case they encountered challenges while working on the case study, they consulted the teaching assistants. After submission, the professor reviewed the case together with the students. In its new, blended learning format, the professor introduces concepts relevant to the case study whereupon the students, in the course of the following week, work entirely using the OILE on their own time and at their own pace. After submission, the professor reviews the case together with the students. However, before submission they do not consult with the teaching assistants because that function is served by the OILE.

5.1 Sample assessments on students’ cognitive development
All the students completed the learning tasks of the OILE. More than 50 percent of the students worked through the case study twice, merely on their own initiative. More than 10 percent of the students worked through the OILE three times or more. This paper presents results only from the students’ first time efforts.

Students’ cognitive domains of learning have been assessed based on the process log information. Based on this information, the students’ learning paths have been drawn using GraphViz software (http://graphs.grevian.org/graph). Figure 3 portrays one student’s learning path while the student was performing a sequence of tasks associated with problem identification and different stages of model building. The green lines represent progress in performance and the red lines represent movement to remedial questions. This student has struggled to perform well in problem identification tasks. Consequently, the student received more support while solving these tasks. Possibly, as a consequence, the student’s performance improved to a significantly higher level of performance while addressing the subsequent model-building tasks.

To further demonstrate the support offered by the OILE and the resulting progress the students demonstrated, an instance of learning is presented below.

To test the students’ understanding of stocks (accumulators) and flows (those that increase/drain the stocks), they were asked to address two cases (Cases 1 and 2) that require graphical integration. Each case has three questions organized in three levels. The questions under Case 1 focus on concepts associated with the maximum level of a stock and the condition that lead to the maximum level. Whereas questions under Case 2 focus on the minimum level of a stock and the condition that lead to the minimum level.

Two questions, Q2.3 and Q2.4 were at the first level of Cases 1 and 2, respectively. These questions are similar to the “department store task” (see Sterman, 2002; Cronin and Gonzalez, 2007). In the department store task, Sterman (2002) gave students a time series graph that shows the number of people that are entering and leaving a department store over a period of 30 min. The students were asked four question. Two of them asked the students to determine when the most and least number of people were in the department store. They turned out to be very difficult questions to answer correctly (Sterman, 2002). Cronin and Gonzalez (2007) also found the same result in a separate study.

In the cases discussed in this paper, the student were asked the two questions (Q2.3 and Q2.4) based on Figure 4. The figure shows the flow of bicycles in to and out of Mr. Wang’s repair shop over the first 80 days. OrderRate refers to the number of bicycles that arrive at Mr. Wang’s shop for repair per day and RepairRate refers to the number of bicycles that are repaired and delivered back to customers per day.

In the first level of Case 1, the students were asked to determine when the number of unrepaired bicycles in the repair shop reaches a maximum during the period in question (see the screenshot of the question in Figure A2). Students were provided with four alternatives and they had to choose one as their answer. Those students who failed to choose the correct answer were asked to branch to another question under Case 1 (level 2 question). This branched question
asked students to determine the condition under which the backlog (physical store of unrepaird bicycles) would be at its highest level (see Figure A3). The main argument for providing such branched question is that if the students understand the condition under which the backlog would be at its highest level, they will be able to answer the question at level 1 when they redo it. Again, the students had four alternatives to choose from.

Those students who failed to respond correctly in the branched question were asked to branch further to another question, which is the lowest in the hierarchy of Case 1 (level 3). At this third level, the students were presented with a metaphor of a bathtub and were asked to distinguish the conditions under which the bathtub’s level would increase, decrease and remain the same (see Figure A4). Here again the students had four alternatives to choose from.

Those students who responded correctly at this third level received the reasons for why their answer was correct and why the other alternatives were incorrect (see Figures A5 and A6). Consequently, these students were asked to move up to a higher level and retry the question they originally failed to answer correctly.

Those students who failed to respond correctly at the third level had two more chances to retry the question. Every time they failed to respond correctly, they each received feedback that informed them why their response was incorrect.

Those students who failed to respond correctly in their third attempt received feedback that informed them of the correct answer and explained for each alternative, why that alternative is a correct/incorrect answer (see Figure A7). Subsequently, the students were asked to move up to a higher level and retry the questions they originally failed to answer.

All the students who had worked on the first question of Case 1, whether they got it right on their first attempt or failed and went down and up the branching questions, moved to the first level of Case 2. Case 2 is an analogous situation of Case 1. But, it challenged the students to determine when the number of unrepaired bicycles in the repair shop reaches a minimum in the first level question (Q2.4) and asked them to recognize the condition that lead to the lowest level of the backlog in the subsequent branching question. The figures and the branching techniques used in Case 2 were similar to Case 1. If the students had learned from the questions and the support they received in Case 1, it would be legit to expect them performance better in Case 2. Research, however, shows that working on the questions without having any intermediate support/feedback has not brought any
change on the students’ performance on the analogous questions (Sterman, 2002; Cronin and Gonzalez, 2007; Cronin et al., 2009).

Table II summarizes the distribution of the students who responded correctly to the two first level questions of Cases 1 and 2 and to the subsequent branching questions. In the OILE, a numbering system was used to differentiate the branched questions from the first level question. Hence, Q2.3.1 and Q2.4.1 represent the second level, and Q2.3.1.1 and Q2.4.1.1 the third level of Case 1 and Case 2, respectively.

Of the 57 students who had used the OILE, only 28 (49 percent) students determined correctly the time at which the highest number of unrepaired bicycles were in the repair shop. Of the same 57 students, the number of students who subsequently determined correctly the time at which the lowest number of unrepaired bicycles were in the shop was significantly larger (39 (68 percent)) compared those who correctly responded at the first level of Case 1.

The table also shows that of those who failed to respond correctly at the first level (51 percent in Case 1 and 32 percent in Case 2), fewer students failed to answer the branching questions compared to the higher percentage of students who failed at the first level. This is true for both Cases.

The result observed in the first question of Case 1 (Q2.3) is similar to the findings of other studies such as Sterman (2002) and Cronin and Gonzalez (2007). However, the 19 percent increase in the students’ performance, from 49 percent in the first question of Case 1 (Q2.3) to 68 percent in the first question of Case 2 (Q2.4), and the relatively better performance observed in the branching questions of both cases were not seen in other studies. This paper argues that such an increase in performance may be attributed to the support (scaffold) that students received using the OILE, i.e. the adequacy of the various levels of support provided by the OILE.

5.2 Assessment on students’ affective domains

Two questionnaires (survey research method, Fowler, 2014) based on prior research (Taylor-Powell and Renner, 2009; Maor and Fraser, 2005; Berkeley Center for Teaching & Learning, n.d.) were distributed among the students to assess the affective aspects of learning. The questionnaires were designed to answer RQ3.

The first questionnaire was administered as soon as the students had completed Tasks 1 and 2 of the learning material, whereas the second was administered after they had completed Tasks 3–5. In total, 53 students out of 57 (93 percent) responded to the two questionnaires. A total of 38 questions were administered through the two questionnaires. These 38 questions are summarized in six major categories as shown in Table III. A detailed account of the students’ response to questionnaire 1 and 2 is provided in Table AII.

The results are as follows:

- More than 75 percent of the students believe that the OILE has a clear user interface and that it is easy to navigate. They also indicate that the texts used in the OILE are not too long to hinder their learning.

<table>
<thead>
<tr>
<th>Question hierarchy</th>
<th>Case 1 (Maximum level)</th>
<th>Case 2 (Minimum level)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1</td>
<td>Q2.3 (49% of total)</td>
<td>Q2.4 (68% of total)</td>
</tr>
<tr>
<td>Level 2</td>
<td>Q2.3.1 (62%* = 32% of total)</td>
<td>Q2.4.1 (61%** = 19% of total)</td>
</tr>
<tr>
<td>Level 3</td>
<td>Q2.3.1.1 (64%** = 12% of total)</td>
<td>Q2.4.1.1 (100%** = 13% of total)</td>
</tr>
</tbody>
</table>

Notes: n = 57. Level 1: questions that asked students to determine when the highest (Q2.3)/lowest (Q2.4) number of unrepaired bicycles were in the shop; Levels 2 and 3: branching questions that support those students who failed to respond correctly to Level 1 questions. *Indicates percent of correct responses from those incorrectly responded at Level 1; **indicates percent of correct responses from those incorrectly responded at Level 2.
More than 77 percent of the students believe that the content is appropriate to their level and they felt they had acquired knowledge in a step-by-step manner.

More than 64 percent of the students believe that they had read and learned from the feedback offered.

More than 80 percent of the students believe they had effectively applied their knowledge and skills acquired in a previous course while working in the OILE.

More than 77 percent of the students claimed that they had understood well the learning material and felt they were ready to move to the next challenge of the course.

More than 84 percent of the students recommended that other system dynamics students, who are at the same level as them, make use of the OILE.

A Wilcoxon Signed-Ranks test was conducted to evaluate whether students showed greater satisfaction in questionnaire 2 compared to questionnaire 1 (see Table IV). The results indicated statistically significant differences for question 3 (attitude toward the feedback offered), $Z = -2.71, p = 0.007$ and for question 5 (belief about their learning), $Z = -2.00, p = 0.046$. The means of the ranks in favor of the OILE in questionnaire 2 were $Mdn = 6.00$.

Table III.

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Questions</th>
<th>Strongly disagree (%)</th>
<th>Disagree (%)</th>
<th>Neither agree or disagree (%)</th>
<th>Agree (%)</th>
<th>Strongly agree (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Experience with the user interface of the OILE</td>
<td>1</td>
<td>11</td>
<td>13</td>
<td>57</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>It has clear interface</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>It is easy to navigate through the OILE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>The OILE does not have unnecessary long texts that hinder my learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Attitude toward the content of the learning and its organization</td>
<td>1</td>
<td>5</td>
<td>17</td>
<td>54</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>It is appropriate to students of my level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I have learned from the tasks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>It helped me learn step-by-step</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Attitude toward the feedback offered</td>
<td>3</td>
<td>18</td>
<td>15</td>
<td>53</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>I have read all the feedback</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I have learned from the feedback</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Attitude toward application of the knowledge and skills they acquired in a previous course</td>
<td>1</td>
<td>6</td>
<td>13</td>
<td>68</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>The OILE gave me the opportunity to practice the skills I gained during a previous course</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Belief about their learning</td>
<td>1</td>
<td>4</td>
<td>18</td>
<td>55</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>I have understood the objective of the case study</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I have understood the main problem in the case study</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I have gotten deeper insight about the main concepts of the case</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I am ready to embark on the next challenge of the course</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Regarding future use of the OILE</td>
<td>1</td>
<td>0</td>
<td>15</td>
<td>46</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>I recommend other system dynamics students of my level to make use the OILE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: $n = 57$. Response rate $53/57 = 93$ percent
(for question 3) and Mdn = 2.50 (for question 5), while the means of the ranks in favor of the OILE in questionnaire 1 were Mdn = 6.00 (for question 3) and Mdn = 0.00 (for question 5). For the other four questions, the Wilcoxon Signed-Ranks test indicated that there were no statistically significant differences in the satisfaction level of the students between the two questionnaires.

Here, it is important to highlight that students had developed a more favorable attitude toward the importance of the feedback they received through the OILE and their attitude had grown more favorably as they progressed through the learning materials. Similarly, their beliefs about their learning became stronger as they progressed through the learning materials. Furthermore, from the analysis of the Wilcoxon test, it can be inferred that the students did not find significant differences across the interfaces of the various OILEs; this signifies the consistency in the design of the OILE.

Based on the overall analyses of the two questionnaires, the paper concludes that students have had a positive experience and developed a friendly attitude toward the OILE.

### 5.3 Intervention strategies and limitation of the study

#### 5.3.1 Intervention strategies

During the design of the OILE, three intervention strategies (immediate, intermediate and long term) were designed to tackle challenges students might face while using the OILE.

The immediate intervention is one that is offered through the OILE and it is provided when students struggle to perform well. It is offered in the form of support as described in section 4.4.2.

The intermediate intervention is one that follows preliminary assessment of the learning analytics. One of the most notable intermediate interventions made was the introduction of a new case study following the preliminary analysis of the OILE’s first time use (during the year 2017). The new case study has a concept similar to the one used to design the Mr. Wang OILE.
It has been designed to reinforce the concepts addressed with the support of the OILE and also to access the transferability of the knowledge and skills gained while using the OILE. It has been designed using a paper and pencil format and was administered to the 2018 system dynamics master program cohort immediately after they had been exposed to the Mr. Wang OILE. Currently, the findings from this new case study and from the OILE experiment are undergoing detailed analyses to be published at a later stage.

The long-term intervention is one that aims at improving the quality of the OILE based on feedback from the questionnaires and analyses of the learning analytics. Results from this intervention will also be communicated in future publications.

5.3.2 Limitation of the study. This paper is delimited to the provision of support to individual students during their study in and about CDS. However, future studies need to document how to foster collaborative learning in and about CDS while accounting for individual students need. Most of the existing platforms that support collaborative learning in and about CDS focus on the dynamics of the groups’ interaction without offering detailed account to the individual students need.

Another major limitation of the study is its inability to generate reports automatically that are easy to read and to interpret. Currently, the students’ data was collected first in the form of CSV files and then manually converted into spreadsheets before the data was coded into learning paths with the help of the GraphViz software (http://graphs.grevian.org/graph). With current advancement in artificial intelligence, the automatic generation of such reports should be possible soon.

A third limitation of the study that is worth of mentioning here is the diversity of the educational media used in the OILE. The study is limited to the use of texts, graphs and simulations. However, in future, the choice of educational media need to be broadened, particular, the inclusion of audios and videos should be considered. The inclusion of such medias could potential increase the learners’ active engagement with the OILE.

With regard to students’ assessment, the current study relied heavily on the use of MCQ and on OEQ that ask students to estimate the over time development of variables that have significant impact on the Mr. Wang’s problem formulation. However, future work should consider diverse assessment techniques, such as questions that address the students’ comprehension, reflective questions and essays that encourage students to describe and respond to the learning tasks using their own words.

Despite the limitations mentioned above, the study showed stronger evidence in the students’ cognitive as well as affective domains of learning. This is evidence of the effectiveness of the proposed design framework aimed at facilitating students’ learning in and about complex, dynamics systems. There is a strong belief that, if properly applied, this five-step design framework for personalized and adaptive learning, together with the three key features of the OILE, would play a significant role in the students’ learning in and about CDS.

6. Conclusion
Research shows the world is facing a wide range of increasingly complex, dynamic problems in both the public and private sectors; climate change, unemployment, health problems, famine, migration, supply-chain problems, etc., create challenges for private and public organizations (OECD, 2017). The problems we face often have a dynamic nature and originate from systems that cause the problem behavior to develop over time. This dynamic development originates from the internal structure of the system (Diehl and Sterman, 1995; Davidsen, 1996). Research demonstrates that we are not well prepared to meet the challenges presented to us in the form of dynamic complexity, neither mentally nor in the form of theories, methods, techniques and tools (Davidsen, 1996; Spector and Anderson, 2000; Jonassen, 2000; Ifenthaler and Eseryel, 2013; Van Merriënboer and Kirschner, 2017).
The objective of this project is to support learning in and about complex, dynamic systems by developing effective instructional methods, techniques and tools; so that students can develop deep intuitions about complex, dynamic systems and an ability to reveal quick fixes that ignore real world complexity (OECD, 2017; Sterman, 2011). For this purpose, the creation of a personalized and adaptive OILE is proposed.

This paper demonstrates how the OILE supports individual learners in their study of CDS. The OILE has been developed based on six models of instructional design that follow a holistic instructional design principle; 4 C/ID, first principles of instruction, CLE, TCI, cognitive apprenticeship and elaboration theory. The OILE has the following three characteristics:

(1) It presents an authentic, complex dynamic problem that the learner should address in its entirety. It then proceeds to allow learners to progress through a sequence of learning tasks from easy to complex.

(2) When solving the problem, learners interact with the OILE. After completion of each learning task, the OILE provides learners with supportive information based on their individual performance. The support fades away as learners gain expertise.

(3) It tracks and collects information on students’ progress and generates learning analytics, which are used to assess students’ learning.

The progress log information was used to assess the cognitive domains of students’ learning, whereas the affective domains of students’ learning were assessed using two questionnaires. Sample results from the progress log of a student demonstrate that the OILE has facilitated the students’ cognitive development. Analyses of the questionnaires show that students firmly believe that they have been through an effective learning experience while working within the OILE.

The literature makes the case for the importance of and difficulty in comprehending CDS. In the study, supportive evidence from the students’ progress in the cognitive domain, as well as their confirmative response in the affective domain, allow the paper to conclude that the use of personalized and adaptive OILE to support learning in and about complex, dynamic systems is promising.

References


Appendix

A.1 Sample time series graph

Figure A1 shows sample time series graph drawn by students. The students were asked to estimate the development of a Backlog (an accumulator) over 150 time units. Initially, the Backlog had 1,000 units and one inflow that initially set to 1,000 units per day. At time 10, the flow rate stepped up by 100 units over time. The actual development and students' estimations are shown in the graph.

![Sample time series graph](image)

**Figure A1.** Sample time series graph: students’ estimate for the development of a backlog over a period of 150 days.

**Notes:** Initial backlog = 1,000 units, Inflow = 1,000 + STEP (100, 10); the black dotted line is the actual development of the backlog; the thin colored lines are students’ estimations.
and remained there for the duration of the simulation:

\[
\text{Inflow} = 1,000 + \text{STEP}(100, 10).
\]

Each line in the time series graph represents one student’s response. To keep the students’ anonymity, personal information was deleted from the time series graph (Figure A1).

A.2 Sample csv files
Table A1 shows sample csv files that have been converted into spreadsheet data. The spreadsheet displays three students learning analytics. To keep the students’ anonymity, personal information was deleted from the spreadsheet. The spreadsheet shows the first 31 data points for each student, which include the different pages the students had visited, the time the students had arrived on a specific page and the amount of time the students had spent on the page. During the analysis, the pages have been coded as either welcome page, task description page, item description page or feedback page. These coded pages served as a base for drawing the students’ learning paths.
Table AI.
Sample csv file of three students converted into spreadsheet data.

<table>
<thead>
<tr>
<th>Page number</th>
<th>Student A</th>
<th>Time spent (seconds)</th>
<th>Page number</th>
<th>Student B</th>
<th>Time spent (seconds)</th>
<th>Page number</th>
<th>Student C</th>
<th>Time spent (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>October 9, 2017</td>
<td>1</td>
<td>1</td>
<td>October 9, 2017</td>
<td>1</td>
<td>October 9, 2017</td>
<td>35.093</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>12:51:03+00</td>
<td></td>
<td>2</td>
<td>12:56:42+00</td>
<td></td>
<td>12:54:30+00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>12:54:13+00</td>
<td></td>
<td>3</td>
<td>12:53:33+00</td>
<td></td>
<td>12:55:05+00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>12:54:20+00</td>
<td></td>
<td>4</td>
<td>12:55:34+00</td>
<td></td>
<td>12:55:58+00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>12:55:04+00</td>
<td></td>
<td>5</td>
<td>12:58:24+00</td>
<td></td>
<td>12:57:54+00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>12:55:25+00</td>
<td></td>
<td>6</td>
<td>12:58:27+00</td>
<td></td>
<td>12:58:55+00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>12:56:14+00</td>
<td></td>
<td>7</td>
<td>12:58:32+00</td>
<td></td>
<td>13:00:14+00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>12:56:44+00</td>
<td></td>
<td>8</td>
<td>12:58:40+00</td>
<td></td>
<td>13:03:52+00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>12:58:56+00</td>
<td></td>
<td>9</td>
<td>12:58:45+00</td>
<td></td>
<td>13:04:31+00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>12:56:19+00</td>
<td></td>
<td>10</td>
<td>12:58:34+00</td>
<td></td>
<td>13:14:41+00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>13:04:20+00</td>
<td></td>
<td>11</td>
<td>12:56:17+00</td>
<td></td>
<td>13:14:53+00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>12:56:17+00</td>
<td></td>
<td>12</td>
<td>12:56:27+00</td>
<td></td>
<td>13:15:30+00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>12:56:30+00</td>
<td></td>
<td>13</td>
<td>12:56:30+00</td>
<td></td>
<td>13:15:34+00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>12:58:29+00</td>
<td></td>
<td>14</td>
<td>12:58:30+00</td>
<td></td>
<td>13:18:36+00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>12:58:48+00</td>
<td></td>
<td>15</td>
<td>13:09:30+00</td>
<td></td>
<td>13:18:53+00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>13:09:30+00</td>
<td></td>
<td>16</td>
<td>13:09:30+00</td>
<td></td>
<td>13:18:53+00</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Page number</th>
<th>Student A</th>
<th>Time spent (seconds)</th>
<th>Page number</th>
<th>Student B</th>
<th>Time spent (seconds)</th>
<th>Page number</th>
<th>Student C</th>
<th>Time spent (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>October 9, 2017 12:58:55+00</td>
<td>64.003</td>
<td>23</td>
<td>October 9, 2017 13:01:27+00</td>
<td>80.849</td>
<td>51</td>
<td>October 9, 2017 13:21:12+00</td>
<td>3.304</td>
</tr>
<tr>
<td>7</td>
<td>October 9, 2017 12:59:59+00</td>
<td>33.669</td>
<td>35</td>
<td>October 9, 2017 13:11:58+00</td>
<td>60.323</td>
<td>53</td>
<td>October 9, 2017 13:21:16+00</td>
<td>10.238</td>
</tr>
<tr>
<td>8</td>
<td>October 9, 2017 13:00:33+00</td>
<td>20.44</td>
<td>36</td>
<td>October 9, 2017 13:12:58+00</td>
<td>39.526</td>
<td>54</td>
<td>October 9, 2017 13:21:20+00</td>
<td>39.214</td>
</tr>
<tr>
<td>9</td>
<td>October 9, 2017 13:00:53+00</td>
<td>172.477</td>
<td>48</td>
<td>October 9, 2017 13:13:37+00</td>
<td>97.889</td>
<td>61</td>
<td>October 9, 2017 13:22:05+00</td>
<td>118.888</td>
</tr>
<tr>
<td>10</td>
<td>October 9, 2017 13:04:34+00</td>
<td>24.513</td>
<td>51</td>
<td>October 9, 2017 13:15:15+00</td>
<td>4.678</td>
<td>62</td>
<td>October 9, 2017 13:24:04+00</td>
<td>3.331</td>
</tr>
<tr>
<td>22</td>
<td>October 9, 2017 13:06:41+00</td>
<td>125.509</td>
<td>53</td>
<td>October 9, 2017 13:15:20+00</td>
<td>3.448</td>
<td>74</td>
<td>October 9, 2017 13:24:07+00</td>
<td>14.358</td>
</tr>
<tr>
<td>26</td>
<td>October 9, 2017 13:07:12+00</td>
<td>3.77</td>
<td>54</td>
<td>October 9, 2017 13:15:23+00</td>
<td>16.659</td>
<td>75</td>
<td>October 9, 2017 13:24:22+00</td>
<td>62.341</td>
</tr>
<tr>
<td>27</td>
<td>October 9, 2017 13:08:05+00</td>
<td>3.419</td>
<td>61</td>
<td>October 9, 2017 13:15:30+00</td>
<td>37.56</td>
<td>76</td>
<td>October 9, 2017 13:25:24+00</td>
<td>86.649</td>
</tr>
<tr>
<td>32</td>
<td>October 9, 2017 13:09:25+00</td>
<td>4.735</td>
<td>74</td>
<td>October 9, 2017 13:16:21+00</td>
<td>15.282</td>
<td>78</td>
<td>October 9, 2017 13:27:50+00</td>
<td>129.099</td>
</tr>
<tr>
<td>33</td>
<td>October 9, 2017 13:09:25+00</td>
<td>16.884</td>
<td>75</td>
<td>October 9, 2017 13:16:36+00</td>
<td>65.032</td>
<td>79</td>
<td>October 9, 2017 13:29:50+00</td>
<td>32.534</td>
</tr>
<tr>
<td>35</td>
<td>October 9, 2017 13:09:25+00</td>
<td>29.947</td>
<td>74</td>
<td>October 9, 2017 13:17:41+00</td>
<td>2.899</td>
<td>80</td>
<td>October 9, 2017 13:30:32+00</td>
<td>5.975</td>
</tr>
<tr>
<td>36</td>
<td>October 9, 2017 13:09:25+00</td>
<td>2.022</td>
<td>75</td>
<td>October 9, 2017 13:17:44+00</td>
<td>19.42</td>
<td>114</td>
<td>October 9, 2017 13:30:38+00</td>
<td>16.991</td>
</tr>
<tr>
<td>48</td>
<td>October 9, 2017 13:09:25+00</td>
<td>74.982</td>
<td>76</td>
<td>October 9, 2017 13:18:03+00</td>
<td>96.881</td>
<td>115</td>
<td>October 9, 2017 13:30:55+00</td>
<td>3.85</td>
</tr>
<tr>
<td>50</td>
<td>October 9, 2017 13:09:25+00</td>
<td>1.714</td>
<td>77</td>
<td>October 9, 2017 13:19:40+00</td>
<td>192.198</td>
<td>149</td>
<td>October 9, 2017 13:30:58+00</td>
<td>29.314</td>
</tr>
<tr>
<td>53</td>
<td>October 9, 2017 13:09:25+00</td>
<td>1.637</td>
<td>78</td>
<td>October 9, 2017 13:22:52+00</td>
<td>87.226</td>
<td>150</td>
<td>October 9, 2017 13:31:28+00</td>
<td>33.831</td>
</tr>
</tbody>
</table>

**Note:** Personal information such as student name and e-mail address that help to identify individual students are removed from the sample csv file.
<table>
<thead>
<tr>
<th>S. No.</th>
<th>Questions</th>
<th>Questionnaire</th>
<th>Strongly disagree (%)</th>
<th>Disagree (%)</th>
<th>Neither agree or disagree (%)</th>
<th>Agree (%)</th>
<th>Strongly agree (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Experience with the user interface of the OILE</td>
<td>Q1</td>
<td>2</td>
<td>11</td>
<td>11</td>
<td>57</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>It has clear interface</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>It is easy to navigate through the OILE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>The OILE does not have unnecessary long texts that hinder my learning</td>
<td>Q2</td>
<td>0</td>
<td>11</td>
<td>15</td>
<td>57</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td></td>
<td>1</td>
<td>11</td>
<td>13</td>
<td>57</td>
<td>18</td>
</tr>
<tr>
<td>2</td>
<td>Attitude toward the content of the learning and its organization</td>
<td>Q1</td>
<td>2</td>
<td>6</td>
<td>17</td>
<td>51</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>It is appropriate to students of my level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I have learned from the tasks</td>
<td>Q2</td>
<td>0</td>
<td>4</td>
<td>17</td>
<td>56</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>It helped me learn step-by-step</td>
<td>Average</td>
<td>1</td>
<td>5</td>
<td>17</td>
<td>54</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td>Attitude toward the feedback offered</td>
<td>Q1</td>
<td>2</td>
<td>25</td>
<td>13</td>
<td>49</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>I have read all the feedback</td>
<td>Q2</td>
<td>4</td>
<td>11</td>
<td>17</td>
<td>57</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>I have learned from the feedback</td>
<td>Average</td>
<td>3</td>
<td>18</td>
<td>15</td>
<td>53</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>Attitude toward application of the knowledge and skills they acquired in</td>
<td>Q1</td>
<td>0</td>
<td>6</td>
<td>13</td>
<td>70</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>a previous course</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>The OILE gave me the opportunity to practice the skills</td>
<td>Q2</td>
<td>2</td>
<td>6</td>
<td>13</td>
<td>66</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>I gained during a previous course</td>
<td>Average</td>
<td>1</td>
<td>6</td>
<td>13</td>
<td>68</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>Belief about their learning</td>
<td>Q1</td>
<td>2</td>
<td>4</td>
<td>19</td>
<td>53</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>I have understood the objective of the case study</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I have understood the main problem in the case study</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I have gotten deeper insight about the main concepts of the case</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I am ready to embark on the next challenge of the course</td>
<td>Q2</td>
<td>0</td>
<td>4</td>
<td>17</td>
<td>57</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td></td>
<td>1</td>
<td>4</td>
<td>18</td>
<td>55</td>
<td>22</td>
</tr>
<tr>
<td>6</td>
<td>Regarding future use of the OILE</td>
<td>Q1</td>
<td>0</td>
<td>0</td>
<td>19</td>
<td>45</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>I recommend other system dynamics students of my level to make use the</td>
<td>Q2</td>
<td>2</td>
<td>0</td>
<td>11</td>
<td>47</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>OILE</td>
<td>Average</td>
<td>1</td>
<td>0</td>
<td>15</td>
<td>46</td>
<td>38</td>
</tr>
</tbody>
</table>

Notes: $n=57$. Response rate $53/57=93\text{ percent}$. Q1: questionnaire 1, Q2: questionnaire 2, Average: average of Q1 and Q2
A.4 Sample learning instances
Sample questions that tested the students understanding of stocks (accumulators) and flows (those that increase/drain the stocks).

Figure A2 a screenshot of question Q2.3 that asked students to determine when the number of unrepaired bicycles in the Mr. Wang repair shop reaches a maximum during the first 80 days.

Those students who failed to respond question Q2.3 in a correct way were asked to branch to question Q2.3.1 (Figure A3).

Those students who failed to respond correctly to question Q2.3.1 were asked to further branch to question Q2.3.1.1 (Figure A4).

Figure A2. Screenshot of the first level question of Case 1 (Q2.3)

Figure A3. Screenshot of a branching question under Case 1 (level 2)
Students who responded correctly to question Q2.3 were offered a feedback (Figure A5) that informed them the reason why their answer was correct and why the other alternatives were incorrect.

Those students who responded correctly to question Q2.3.1 were offered a feedback (Figure A6) that informed them the reason why their answer was correct and why the other alternatives were wrong.
Those students who attempted question Q2.3.1.1 three times and failed to respond correctly were offered a feedback (Figure A7) that informed them of the correct answer and explained for each alternative, why that alternative is a correct/incorrect answer.

Figure A7. Screenshot of a feedback offered to those who attempted Q2.3.1.1 three times and failed

Corresponding author
Aklilu Tilahun Tadesse can be contacted at: aklihutt@gmail.com
What do first-year students need? Digital badges for academic support to enhance student retention

Dana-Kristin Mah
University of Potsdam, Potsdam, Germany, and
Dirk Ifenthaler
Curtin University, Bentley, Australia

Abstract
Purpose – The purpose of this paper is to analyse data on first-year students’ needs regarding academic support services and reasons for their intention to leave the institution prior to degree completion. On the basis of the findings, a digital badge outline is proposed which could contribute to improved communication of academic requirements in order to help students to better adapt to higher education demands. Digital badges might also serve as an indicator for students’ needing additional academic support services.

Design/methodology/approach – An online-questionnaire was conducted with 730 first-year students at a German university. Participants’ responses to open-ended questions were coded and categorised. On the basis on these findings, an outline for a digital badge programme is proposed.

Findings – Participants seek the most institutional support regarding research skills and organisational aspects. Main reasons for participants’ intention to withdraw from the institution include difficulties with their programme choice.

Practical implications – These findings may enable higher education institutions to provide targeted support services that meet first-year students’ needs. On the basis of the findings, higher education institutions can create digital badge programmes, which may improve communication of academic requirements and may also serve as a platform for a staff-student conversation about expectations and demands for a successful first-year experience. Besides, further research and discussion may address using digital badges for learning analytics algorithms to even better identify students’ strengths and needs for targeted academic support services and enhanced student success in higher education.

Originality/value – Little is known about first-year students’ needs for institutional support and reasons for thinking about dropout in Germany. Understanding the student perspective is crucial for enhancing student retention. Digital badges are an emerging educational technology in higher education and they have the potential to target academic requirements, which may guide first-year students and help them to better adjust to universities’ demands.

Keywords Higher education, Retention, Learning analytics, First-year students

Paper type Research paper

1. Introduction
Supporting first-year students in the transition from secondary to tertiary education is an important issue for higher education institutions in order to enhance student retention. Personal and adaptive academic support services should be offered to meet students’ individual needs (Schumacher and Ifenthaler, 2018a). This is particularly important with regard to students’ diversity (Bosse, 2015; Reason et al., 2006). Many higher education institutions already offer support services such as orientation seminars, freshman courses, and mentoring programmes (Clark and Cundiff, 2011; Tinto, 2012). One-third of the students, however, still withdraw from the institution prior to degree completion (OECD, 2013), highlighting that further research is needed. In Germany, the most recent study on student dropout reports an overall dropout rate of 32 per cent in the Bachelor programme (Heublein et al., 2017). The highest dropout rates are found for mathematics and natural sciences (39 per cent) and engineering (32 per cent) and the lowest for arts and cultural studies (23 per cent).
The first year of higher education has been identified as a crucial period with most of the students decide to dropout within or shortly after their first year (Heublein et al., 2017; Tinto, 1975). Reasons for student dropout prior to degree completion are many and can affect one another (Heublein, 2014). Challenges include academic transition, a social transition, and meeting of expectations and perceptions (Kantanis, 2000; Smith and Wertlieb, 2005). Furthermore, academic competencies, which focus on interdisciplinary generic skills (Clanchy and Ballard, 1995), may influence student retention. For instance, Mah and Ifenthaler (2018) found, that first-year students’ perceptions of low academic support in developing research skills for their studies is related to their intention to leave the institution. Furthermore, first-year students are often unsure of what is expected of them in academic terms while not coping with academic requirements is an important factor for deciding to discontinue higher education (Thomas, 2002; Yorke and Longden, 2008). Academic staff, however, often expect first-year students to be academically prepared for higher education and to have the capacity to cope with higher education demands on the basis of their prior school education (Barrie, 2007). Thus, clear communication of expectations is essential for a successful transition to higher education studies and approaches have to be considered accordingly.

Transparency is one main objective of digital badges. Digital badges in education symbolise achievements, knowledge, skills, and competencies in various educational contexts. They can play four main roles in education: motivation, recognition of learning, signalling of achievements and capturing of learning paths, all of which have the potential to contribute to student retention in higher education (Mah et al., 2016).

As more and more educational data are available, learning analytics use static and dynamic data about learners and learning environments to assess, elicit, and analyse it for real-time modelling, prediction and optimising of learning processes (Ifenthaler, 2015). Many studies show a positive impact on learning analytics regarding student retention in higher education by identifying student at risk, providing personalised feedback and guidance to academic support services (Ifenthaler et al., 2019; Pistilli, 2017; Sclater et al., 2016).

The purpose of this case study is to provide insight into first-year students’ needs regarding academic support services; second, their main reasons for their intention to dropout from higher education. Moreover, on the basis of these findings, third, an outline for a digital badge to be developed as a means of improving the communication of requirements is proposed. This study is an attempt to link data on first-year students’ needs regarding academic support services and the concept of digital badges as a means of meeting those needs in higher education.

2. Support services in higher education

Institutional support services are crucial for students, especially for first-year students (Yorke, 2000). Institutional support services often include summer bridge programmes, first-year seminars and mentoring programmes (Tinto, 2012). However, academic support services differ by institution as well as between national contexts (Padgett et al., 2013). Studies show that not all students are informed about the available support services (Banscherus et al., 2015). Thus, information events, advertisement, networks, and improved communication may be useful to introduce these support services to incoming students (Mah and Ifenthaler, 2018). Furthermore, studies show that many students do not participate in voluntary academic support programmes, even though research shows their effectiveness (Attewell et al., 2006; Schmied and Hänze, 2015). Hence, student support might be more effective as an integrated part of the curriculum aligned in the classroom (Tinto, 2012). Also, support services should be made available face-to-face and online to meet students individual needs.
3. Educational data and digital badges in higher education

Educational data in higher education has been increasing in the past few years. Educational technologies for learning and teaching in higher education such as learning analytics, have become more established and are important drivers in changing learning and teaching environments (Arnold, 2010; Ifenthaler, 2017; Sclater and Mullan, 2017; Sclater et al., 2016). With regard to the use of educational data and the rise of educational technologies for improving learning and teaching, digital badges is a technology, which shows promise when adopted to enhancing student retention (Ifenthaler et al., 2016). Digital badges are symbols of learning achievements, skills and competencies across educational contexts (Gibson et al., 2013). Learners can collect the digital images in their personal badge system, and display them on social media platforms and professional networks such as LinkedIn (Glover, 2016). There are different types of digital badges, for example for achievement, capability, membership, and participation (Belshaw, 2015), and different levels, such as basic knowledge badges, specific skills badges, and advanced knowledge badges (Pöldoja and Laanpere, 2014). Digital badges can have different objectives in education such as enhancing motivation, recognition of learning, signalling of achievements, and capturing of learning paths (Jovanovic and Devedžić, 2015; Mozilla Foundation and Peer 2 Peer University, 2012). Furthermore, digital badges' contribution to students’ first-year experience and student retention has been discussed (Mah et al., 2016; Mozilla Foundation and Peer 2 Peer University, 2012). In this regard, first-year students could feel motivated to achieve digital badges that recognise and verify their learning within the higher education institution, as well as in informal settings and from previous experiences (e.g. job experience). Learners’ achievements could be signalled and learning paths captured, which may provide structure for first-year students in their transition to higher education (Mah et al., 2016). Recent research discusses the potential of digital badges for students’ goal setting which may impact learning performance (Cheng et al., 2018). Successful implementation of digital badges, outlines and digital badge development guidelines have been proposed. For example, the worksheet by Wright and O'Shea (2014) includes information on the badge issuer, badge name and description, target audience, learning outcomes, learning activities, and required evidence and assessment criteria. Mah et al. (2016) present guidelines for designing digital badges in the Passport platform created by Purdue University. This platform guides the developer through three main sections, which are general badge information, the badge image, and the challenge. Furthermore, a synthesis of digital badges and learning analytics may be promising in an effort to enhance student retention in higher education: Interdisciplinary, generic skills can be represented as digital badges, which can be used with learning analytics algorithms to predict student success and to provide learners with personalised feedback and guidance to academic support services (Mah et al., 2016).

Both, quantitative data and qualitative findings may be worth exploring and considered when designing digital badges. This study uses findings from a student questionnaire in order to create an exemplary digital badge outline that may support first-year students’ needs.

4. Research questions, data collection and method

4.1 Research questions

The present case study seeks to explore first-year students’ perspective regarding their need of academic support and reasons for thinking about leaving the institution prior to degree completion. On the basis of the first findings, an exemplary digital badge outline is proposed. The research questions were as follows:

RQ1. In which areas do first-year students require support from their institution?

RQ2. Why are first-year students thinking about leaving the institution?
RQ3. How can a digital badge outline for an exemplary requirement for higher education studies be developed which takes into account aspects such as course outcomes, learning activities, and required evidence and assessment criteria?

4.2 Data collection and sample
Data were collected using an online-questionnaire in the Spring semester 2015 at a German higher education institution. The questionnaire consists of closed-questions, open-ended questions and socio-demographic questions. The sample consists of 730 first-year students with a mean age of 20.1 year (SD = 2.0) with 59.7 per cent female students. The majority of participants was enrolled in the department of economics (52.5 per cent), followed by the school of humanities (17.4 per cent), the school of business informatics and mathematics (11.9 per cent), the department of law (11.2 per cent) and the school of social sciences (7.0 per cent). Most of the participants (39.7 per cent) are assigned to the study programme “Economic and Business Education”, which qualifies to work at a school among other options. The average grade to attend higher education was 2.2 (SD = 0.6, Min. = 1.0, Max. = 3.7), with the scale ranging from 1 = excellent to 6 = very poor).

4.3 Method
In order to answer RQ1 and RQ2, first-year students open-ended responses to the questions “In which areas do you wish/require support from your institution?” and “Why are you currently thinking about leaving the institution?” were analysed and coded using MAXQDA analysis software. Thematic data analysis (Flick, 2014) includes inductive coding of participants’ statements in order to develop a frame of descriptive categories. These main categories focus on the relevant aspects with regard to the research questions. A short description and sample statements are provided as representation of each category. With regard to the findings, an outline for an exemplary digital badge is proposed, which may help to address and support students’ expressed needs in their first year of higher education studies.

5. Results
5.1 Academic support requests
In total, 436 first-year students responded to the open-ended question regarding their academic support requests: “In which areas do you wish/require support from your institution?” Participants’ responses (451 statements) were clustered into nine main categories: organisation, research skills, supervision, learning skills, exam preparation, time management, practical relevance, transparency and other aspects. Table I provides descriptions/aspects of the nine main categories including sample statements and frequencies.

5.2 Reasons for dropout intention
Overall, regarding first-year students’ intention (frequency) to leave the institution, 66.7 per cent responded never, 20.7 per cent rarely, 9.7 per cent occasionally, 2.5 per cent frequently und 0.5 per cent very frequently. There were no significant differences found regarding socio-demographics such as gender ($U = 61,215.50, p > 0.05$), age [$F(15,714) = 0.951, p > 0.05$], and average grade to attend higher education $F(27,700) = 1.050, p > 0.05$). Also, there were no significant differences found for their father’s [$\chi^2(5) = 3.82, p > 0.05$] and mother’s [$\chi^2(5) = 3.76, p > 0.05$] highest school-leaving qualification. With a focus on the open-ended question “Why are you currently thinking about leaving the institution?” a total of 130 participants responded. Table II presents sample statements and frequencies for the identified four main categories as a result of the inductive coding process (138 statements).
<table>
<thead>
<tr>
<th>Category (Subcategory)</th>
<th>Description/Aspects</th>
<th>Sample statement</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organisation</td>
<td>University life, study organisation, university systems (e.g. LMS)</td>
<td>Information to organise my studies</td>
<td>132</td>
</tr>
<tr>
<td></td>
<td></td>
<td>General introduction to university (campus, online systems)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Study plan</td>
<td></td>
</tr>
<tr>
<td>Research skills</td>
<td>Academic writing, essays, literature search</td>
<td>How do I write academic essays and short reports?</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Research methods</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>How do I find adequate literature?</td>
<td></td>
</tr>
<tr>
<td>Supervision</td>
<td>Availability of academic staff, counselling, tutorials</td>
<td>Better communication with academic staff</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>More contact to academic staff</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Personalised support and feedback</td>
<td></td>
</tr>
<tr>
<td>Learning skills</td>
<td>Learning strategies, dealing with learning content, effective learning methods</td>
<td>Selection of important information</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Learning for exams</td>
<td></td>
</tr>
<tr>
<td>Exam preparation</td>
<td>Mock exams, study material, preparation and support by academic staff</td>
<td>Dealing with a lot of learning content</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>More support regarding exams</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mock exams to get a clue about the exam</td>
<td></td>
</tr>
<tr>
<td>Time management</td>
<td>Time management methods</td>
<td>Access to study materials and literature</td>
<td>26</td>
</tr>
<tr>
<td>Practical relevance</td>
<td>Practical relevance, internships, career counselling</td>
<td>Practical examples for the job</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Job possibilities</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Career counselling</td>
<td></td>
</tr>
<tr>
<td>Transparency</td>
<td>Transparent communication of expectations</td>
<td>Transparent communication of expectations</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Requirements for essays</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>What are general expectations of academic staff?</td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>Financial advice, opening hours library, cafeteria</td>
<td>Financial advice (e.g. state aid)</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Help with finding an accommodation</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Extended library opening hours</td>
<td></td>
</tr>
</tbody>
</table>

**Table I.**

Support requests  

**Notes:** $n = 436$. Multiple statements were possible for the open-ended question

<table>
<thead>
<tr>
<th>Category (Subcategory)</th>
<th>Description/Aspects</th>
<th>Sample statement</th>
<th>Frequency</th>
</tr>
</thead>
</table>
| Difficulty of the study programme | Excessive demands; too difficult, high workload, stress | I'm overstrained  
I've been doing nothing else but studying since the beginning of the programme
It's very exhausting | 68        |
| Wrong choice of study programme | Wrong choice of study programme, no interest                                    | Wrong choice of study  
Boring
It's just not right for me
Too abstract | 36        |
| Mismatch of expectations | Mismatch of expectations, I expected something else | I expected something else  
I'm unsure if university is right for me
Cultural shock
I'm feeling lonely | 7         |
| Others                 | Problem with accommodation, career perspectives, health problems, overall organisation of the programme | I'm afraid of having bad career perspectives  
Prefer to travel
I've already have a vocational training certificate
Chronic disease | 27        |

**Table II.**

Support requests  

**Notes:** $n = 130$. Multiple statements were possible for the open-ended question
5.3 Digital badge outline for research skills

Findings from RQ1 show that participants request the most institutional support regarding organisational aspects and research skills. Results from RQ2 reveal that most of the participants intend to leave the institution because of difficulties with their study programme. They often feel overstrained, stressed and exhausted. Reasons may refer to subject-specific difficulties but may also refer to interdisciplinary requirements, such as research skills, for which participant seek institutional support (RQ1). Moreover, statements indicate a mismatch of expectations as one reason for participants’ intention to withdraw.

On the basis of these findings, an outline for an example of a digital badge with regard to research skills, is proposed. The prerequisite/basic knowledge digital badge focusses on pre-service teacher studies, as the majority of the participants in this study are assigned to a programme that qualifies to teach at schools (Table III).

6. Discussion

The presented case study provides insight into first-year students’ perspective on requested academic support services and reasons for their intention to leave the institution.

With focus on first-year students’ requests of institutional support, most of the participants refer to organisational aspects such as general introduction to university, the campus, and online-systems like LMS and registering for exams. This finding is in line with research that shows that many incoming students need guidance in their first-year of higher education. For instance, using information and support services was identified as an organisational requirement students perceived as critical for transition to higher education (Trautwein and Bosse, 2017). Many higher education instructions, however, already provide support for organisational issues, such as orientation seminars, campus tours and first-year seminars. Academic support services vary by institution, area, and country (Padgett et al., 2013). Furthermore, many students’ are unsure about academic support offerings or do not participate voluntarily in such programmes (Schmied and Hänze, 2015). Furthermore, many students request institutional support regarding the interdisciplinary domain research skills. Research shows that research skills such as academic writing and literature search is usually not taught in schools and thus is especially challenging for first-year students (Tinto, 1993; Wingate, 2006; Author, 2018). Also, participants request support in other interdisciplinary domains such as time management and learning skills. Learning and teaching styles in higher education differ from those in schools and first-year students need to adapt to the new structure very quickly because many academic staff expect incoming students to possess these skills (Barrie, 2007; Mah and Ifenthaler, 2017).

With focus on first-year students’ intention to withdraw, findings are in line with reasons of students reported in international studies and the little available national research. For example, main factors found in studies include the choice of the wrong course, academic unpreparedness, a lack of university support services, and personal issues, such as financial problems, illness, and family circumstances (Brooker et al., 2017; Yorke and Longden, 2008). Also, participants reported a mismatch of expectations with their actual experience in higher education. This finding is also consistent with studies, which highlight the importance of realistic pre-entry expectations for an effective match and success (Crisp et al., 2009; Smith and Wertlieb, 2005). Thus, higher education institutions should define and clearly communicate their academic requirements as soon as possible, for instance by using digital badges. If academic requirements are transparent, first-year students will know what is expected of them and thus develop the academic competencies needed for higher education right from the beginning (Mah and Ifenthaler, 2017, 2018).

Findings from RQ1 and RQ2 revealed that participants assess intuitional support in the transition to higher education as important. Many students, however, are unsure of
Participants in this study highlighted research skills as one crucial requirement of higher education for which they feel unprepared and seek for institutional support. Digital badge outline for research skills is proposed because other studies indicate first-year students' low confidence and expectations for support of this academic requirement (Mah and Ifenthaler, 2018). Overall, research on digital badges in higher education indicates positive effects on academic requirements and wish that guidelines were more transparent communicated.

<table>
<thead>
<tr>
<th>Outline</th>
<th>Basic knowledge badge for research skills in pre-service teacher studies</th>
<th>Design considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Badge issuer</td>
<td>Course instructor for pre-service teachers at a German university, course instructors' e-mail address</td>
<td>Information about the badge issuer (name, institution, e-mail address)</td>
</tr>
<tr>
<td>Badge name and description</td>
<td>“Being academically competent in research skills” Learners will create a digital badge that is focussed on research skills</td>
<td>Research skills was identified as an academic competency in which first-year students require support from the institution (RQ1)</td>
</tr>
<tr>
<td>Target audience description</td>
<td>First-year students in pre-service teacher studies</td>
<td>As the majority of the participants in this case study are assigned to a programme that qualifies to teach at schools, the exemplary digital badge addresses this audience (pre-service teachers)</td>
</tr>
<tr>
<td>Learning outcomes</td>
<td>Students can describe and explain research skills and their key elements in educational contexts Students can demonstrate how to use research skills for different purposes in educational fields Students can develop research questions for educational purposes Students can create an outline for the adequate use of research questions on a given research question in the educational field</td>
<td>Information what learners should be able to demonstrate with regard to research skills As the focus is on the pre-service teachers, learning outcomes are related to educational purposes Important research skills for higher education studies includes for example methodological knowledge, communication, and information seeking (Gilmore and Feldon, 2010; Meerah et al., 2012)</td>
</tr>
<tr>
<td>Learning activities</td>
<td>Students are presented with academic research methods and reflect their use for different purposes in educational fields Students are presented with best-practice examples and guidelines for academic research in educational fields Students discuss how to develop research questions in educational contexts Students engage in discussions about how to interpret the results of a research study for their profession</td>
<td>Information about learning activities, which are linked directly to Learning Outcomes Learning Activities in research skills for pre-service teachers focus on educational contexts Participants' statements are considered, e.g. How do I write academic essays and short reports? How do I find adequate literature? (Table II)</td>
</tr>
<tr>
<td>Required evidence and assessment criteria</td>
<td>Clear written short report, which presents a developed research question, an outline of how to design the study and a discussion of suitable research methods Students are able to discuss the research methods used in a given study in an educational context Students are able to discuss and interpret results of a given research study in an educational context</td>
<td>Information what learners will submit and how it will be assessed The target audience should be able to deal with these requirements to show their basic skills in research skills</td>
</tr>
</tbody>
</table>

Table III. Outline of a basic knowledge badge for the academic competence research skills in pre-service teacher studies

**Note:** Outline following the digital badge worksheet by Wright and O'Shea (2014)
students’ engagement and motivation (Schumacher and Ifenthaler, 2018b), for instance with regard to co-curricular (Coleman, 2017) and hence should be considered to enhance student retention in higher education.

7. Conclusion and further research

Overall, the presented results may have important implications for higher education institutions, such as, encouraging them to provide stronger support services for first-year students to help them adjust to higher education, reacting to reasons for students’ intention to withdraw, defining academic requirements and clearly communicating. Furthermore, the suggested digital badge may encourage higher education institutions to discuss this relatively new educational technology. Until now, digital badges are mostly implemented in English-speaking countries, such as Australia, the USA and the UK (Muilenburg and Berge, 2016). Digital badges may serve as an indicator for students who need academic support. Thus, digital badges may serve as a platform for communication between staff and students about demands and adaptive support services in order to increase student retention (Mah and Ifenthaler, 2017). Research and concepts are also needed that focus on strategies to integrate digital badges and learning analytics into academic support service programmes in a meaningful way (Mah and Ifenthaler, 2018).

Further research should test the proposed digital badge, which is limited to theoretical considerations so far. Additionally, research on the stakeholders’ acceptance, issues of implementation and data privacy of digital badges should be conducted and considered (Berge and Muilenburg, 2016; Glover, 2016). More insights into first-year students’ perspective on institutional support services and their reasons for considering dropping out are necessary, for example, interview studies might help to gain in-depth insights. More research is needed and should include more higher education institutions and also longitudinal studies for validation of the inductive coded categories and for generalisation of the qualitative findings. The synthesis of digital badges and learning analytics (Mah et al., 2016) for a holistic use of educational data should be addressed in further research. Learning analytics shows promise in enhancing student retention, however, higher education institutions may not be prepared for learning analytics projects (Mah and Ifenthaler, 2017; Sclater et al., 2016).

In summary, even more research on educational data (Ifenthaler et al., 2019), meaningful analyses and educational technologies and growing fields such as digital badges and learning analytics (Mah et al., 2019) is required in order to contribute to student retention in higher education.

References


Flick, U. (2014), An Introduction to Qualitative Research, Sage, Los Angeles, CA and London and New Delhi and Singapore and Washington, DC.


**Further reading**


**Corresponding author**

Dirk Ifenthaler can be contacted at: dirk.ifenthaler@curtin.edu.au
Enrollment patterns and students’ risk of academic difficulty

Robert Matthew DeMonbrun
University of Michigan, Ann Arbor, Michigan, USA

Michael Brown
Iowa State University, Ames, Iowa, USA

Stephanie D. Teasley
University of Michigan, Ann Arbor, Michigan, USA

Abstract
Purpose – Experiencing academic difficulty can deter students’ academic momentum, decreasing the speed with which they complete coursework and increasing the odds that they will not persist to a credential. The purpose of this paper is to expand upon an existing framework that investigates students’ academic difficulty in co-enrolled courses by adding additional co-enrollment variables that may influence academic performance in introductory gateway courses.

Design/methodology/approach – This study uses quantile regression to better understand academic difficulty in co-enrolled courses and the impact that students’ co-enrollment patterns may have on their success in focal introductory gateway courses.

Findings – This study revealed significant relationships between student success and co-enrollment patterns, including: the disciplinary alignment of the course with a student’s major, the student’s co-enrollment in other difficult courses and experiencing below average academic performance in a co-enrolled course. Further, impact of these relationships often differed by students’ performance quantile in the focal course.

Practical implications – The results point to factors related to the student and their co-enrolled courses that faculty, academic advisors and curriculum committees can consider as they design general education requirements within and across disciplinary majors.

Originality/value – This approach advances the understanding of how a prescribed curriculum produces interdependent pathways that can promote or deter students’ success through the organization of curricular requirements and student course taking. The paper provides a generalizable methodology that can be used by other universities to investigate curricular pathways that have the potential to reduce student success.

Keywords Quantile regression, Academic performance, Curriculum analytics, Gateway courses, Undergraduate persistence

Paper type Research paper

Introduction
Graduation rates are an increasingly visible metric of an institution’s commitment to undergraduate education and often used by applicants and their parents to judge the potential “value” of enrolling in one school over another. The disruption of students’ academic momentum in college (Adelman, 1999, 2006), typically through experiencing academic difficulty early in their course trajectories, can be detrimental to the progression of their academic careers and their eventual graduation. Academic momentum is defined as the speed with which students’ progress in their studies (Adelman, 1999). Prior research supports the idea that decreasing the speed with which students’ complete their coursework increases the odds that they will not persist to a credential (Adelman, 1999, 2006; Attewell et al., 2012). Academic momentum is also a significant predictor of academic success in the first two years of college (Belfield et al., 2016), retention in undergraduate education (Kondratjeva et al., 2017) and persistence toward science, technology, engineering and math (STEM) degrees (Wang, 2016). Academic momentum may have a sufficient gravitational force of its own that eventually overcomes initial variations in academic preparation (Attewell et al., 2012). Even when other important explanations for differences in achievement, like prior life experiences and academic preparation, are held constant, students with academic momentum are still more likely to achieve and persist (Martin et al., 2013).
The design and organization of the undergraduate curriculum is one of the major barriers to maintaining academic momentum in undergraduate education (Slim et al., 2014a). Complex curricular requirements that result in students taking multiple difficult courses simultaneously have a demonstrable impact on college student success within a course (Brown et al., 2018; DeMonbrun et al., 2018), as well as time to degree and persistence metrics (Heileman et al., 2017; Slim et al., 2014a, b; Wigdahl et al., 2014). When students encounter complex arrangements of their coursework, they have higher odds of experiencing academic difficulty (Brown et al., 2018), which can result in poor course performance in one or more courses. Furthermore, when students fail in early course requirements, either within their major field or as part of their general education, their momentum to degree is slowed.

When curricular complexity deters students’ academic momentum, important college and post-graduate outcomes are potentially impacted. For example, the number of credits a student takes within a semester can alter time to degree across a variety of contexts (Attewell and Monaghan, 2016; Calcagno et al., 2007; Hodara and Rodriguez, 2013), a finding that, while perhaps intuitive on its face, is not sufficiently acknowledged when faculty engage in academic planning (Lattuca and Stark, 2009). As a result, students stall out in progress toward their degree because of narrow, complex and inefficient course pathways (Slim et al., 2014a) – a problem that is compounded in STEM fields (Aldrich, 2015; Heileman et al., 2017).

To understand the relationship between curricular complexity and student performance, in our earlier work, we focused on a gateway engineering course and utilized performance data generated by an Early Warning System (EWS). The EWS provides a weekly categorization of each student’s performance on a course-by-course basis, designating one of three classifications: “Encourage” (green – student performing at or above the course mean), “Explore” (yellow – students performing below the course mean) or “Engage” (red – students in the lowest quartile of performance) (see Krumm et al., 2014).

Using EWS data, we have found several factors that shape students’ potential for academic success. These factors inform the conceptual framework that guides our analysis. First, we found that the discipline of the course – using Biglan’s (1973) typology of disciplinary fields – is significantly related to experiencing academic difficulty in a course, as students were the most likely to experience difficulty in fields like the natural sciences, mathematics, and engineering in comparison to humanities courses (Brown et al., 2016). Second, as the number of co-enrolled “difficult courses” (determined by the percentage of students in the class who entered the “Explore” and “Engage” classifications in each of these courses) increases, so do students’ odds of experiencing academic difficulty in an introductory computer science gateway course (i.e. an introductory core course that most Computer Science Engineering students take as a general requirement; Brown et al., 2018). Finally, students experiencing academic difficulty in one course are much more likely to experience difficulty in multiple courses, thereby producing a “snowball” effect of academic difficulty because of curricular complexity across multiple courses (DeMonbrun et al., 2018).

In this paper, we expand upon the findings in our prior research exploring the role of curricular and students’ academic success by utilizing data from two very different introductory gateway courses. In doing so, we make several significant contributions. First, we observe if patterns of academic difficulty related to coursework in an early major requirement for engineering are also significant for required introductory courses in two other disciplinary pathways: social science (through a general education Economics requirement) and life science (through a general education Biology requirement) majors. Second, we offer insight into how students’ performance in co-enrolled courses might influence their performance in required gateway courses. We hypothesize that students who are performing well in a focal course might be less susceptible to the “snowball” effect of curricular complexity we have observed in our prior work (DeMonbrun et al., 2018). Finally, we offer an analytical method that does not require performance data from EWSs, as has been used in
prior analyses (Brown et al., 2016, 2017, 2018; DeMonbrun and Brown, 2017; DeMonbrun et al., 2018). This allows institutions that do not use these systems to replicate our method using learning analytics data from their own gateway courses. We picked three percentile ranks to test our research questions across different performance levels, specifically the 10th, 50th and 90th percentiles, which also map onto the performance categories determined by the EWS. Thus, we propose the following research questions:

**RQ1.** Are certain co-enrollment patterns related to changes in students’ academic performance in a focal introductory gateway course?

**RQ1a.** What are the associations between the disciplinary typology of a student’s major and their academic performance in the focal course?

**RQ1b.** What are the associations between a student’s enrollment in “difficult courses” and their academic performance in the focal course?

**RQ1c.** What are the associations between the number of co-enrolled courses and a student’s academic performance in the focal course?

**RQ1d.** What are the associations between below average academic performance in a co-enrolled (non-focal) course and a student’s academic performance in the focal course?

**RQ2.** How do these co-enrollment patterns differ by students’ percentile rank – specifically for students in the 10th, 50th and 90th percentiles – in these courses?

## Methodology

### Sample

For this study, we identified two gateway courses at a large, public co-educational research university in the Midwestern USA. These were introductory courses in Biology (Ecology and Evolutionary Biology) and Economics (Microeconomics of Capitalism). Both courses serve as requirements for two of the three largest colleges at the university. Our sample includes 1,785 undergraduate students who enrolled in at least one of the two courses in either the Fall or Winter terms of the 2017–2018 academic year. Demographics for students in the sample are shown, by class, in Table I. Students in the Biology courses were 68 percent female, predominantly White (62–65 percent), and around a quarter first-year students (23–29 percent). Students in the Economics courses were predominantly male (39–41 percent female) and White (58–63 percent), while the percentage of first-year students differed greatly between Fall (17 percent first-year students) and Winter (48 percent first-year students) terms.

<table>
<thead>
<tr>
<th></th>
<th>Ecology and evolutionary biology</th>
<th>Microeconomics of capitalism</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fall (n = 558)</td>
<td>Winter (n = 546)</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Female</td>
<td>382</td>
<td>68.46</td>
</tr>
<tr>
<td>White</td>
<td>363</td>
<td>65.05</td>
</tr>
<tr>
<td>Black</td>
<td>18</td>
<td>3.23</td>
</tr>
<tr>
<td>Hispanic</td>
<td>33</td>
<td>5.91</td>
</tr>
<tr>
<td>Asian</td>
<td>97</td>
<td>17.38</td>
</tr>
<tr>
<td>Multiracial</td>
<td>24</td>
<td>4.30</td>
</tr>
<tr>
<td>Unknown</td>
<td>23</td>
<td>4.12</td>
</tr>
<tr>
<td>International</td>
<td>8</td>
<td>1.43</td>
</tr>
<tr>
<td>First-Year Student</td>
<td>130</td>
<td>23.30</td>
</tr>
</tbody>
</table>

**Note:** n = 1,785
Variables

All gradebook data for the Biology and Economics courses were pulled from the university’s data warehouse – with approval from the university’s institutional review board and ensuring all data were anonymized – to provide a weekly update of each student’s academic performance, which serves as the dependent variable for this analysis. At the end of each week (Saturday), the data warehouse was updated with that week’s course gradebook data, providing a list of earned and possible points for all assignments for each student in the sample during that week. Then, we took the updated points and added them to the existing number of points earned vs points possible from all assignments in prior weeks to create an updated weekly course grade for each student, which was modeled as a percentage of the number of points earned divided by the possible number of points.

The independent variables of interest are the co-enrollment patterns and performance of students in a focal course (Biology or Economics). First, because of the wide variety of academic majors in our sample, we grouped students’ declared majors at the beginning of the academic term using Biglan’s (1973) typology of disciplinary fields. Biglan’s typology classifies fields based on the consensus around core and new knowledge in the field, and the purpose for which new knowledge is being developed (pure theory or application). Additionally, Biglan classifies fields based on whether they focus on organic or inorganic subjects. There are three dimensions in Biglan’s typology: hard/soft (the degree of consensus on core research questions, research methods, and knowledge within the field); pure/applied (new knowledge is produced to further theory or practice applications); and life/non-life (the extent to which living or organic objects are the focal point of the field’s research). Each major has one designation for each dimension. We also created a separate variable for all students who had not yet declared a major program (undeclared).

Second, we developed a measure to determine the number of “difficult courses,” not including the focal course in this analysis, in which each student was co-enrolled during the semester. A difficult course was classified as any course where the average student grade was below 70 percent of the total possible grade. This threshold was chosen because grades below a 70 percent are classified by the university as a D+ and below, which for many students indicates that the course must be retaken in order to progress forward in their major.

Third, to determine the impact of students’ course load (i.e. the number of courses taken simultaneously in a semester), we created a measure for the total number of courses – including the focal course – in which each student was enrolled. Finally, we also created a measure to determine the number of co-enrolled courses where students, themselves, were experiencing below average academic performance using the same < 70 percent threshold, not including the focal course.

We also added controls for various demographic and academic characteristics in each model, including variables for gender, race/ethnicity, citizenship status, first-year students (by credit hours), math placement test (a test scored from 1 to 25 and required of all incoming students at the university to determine proper sequencing in any math-related courses for their major), the student’s high school GPA upon entering college, and their incoming college GPA if the student had taken prior college courses. After initial testing for model fit, we determined that the combination of high school GPA, college GPA and math placement test led to issues with multicollinearity in the models. We tested the strength of these three variables by inserting each variable separately into the model to determine which variable best explained the variation in weekly academic performance and found that the math placement test variable was the best explanatory variable available. Thus, all subsequent analyses report only the association between math placement test and students’ academic performance. We also included time indicators as a measure of the longitudinal nature of this data, and whether or not grade fluctuations were linear or non-linear in nature.
Thus, we included controls for linear, quadratic, and cubic relationships with time, consistent with our prior analyses (DeMonbrun et al., 2018).

We should also note that, while we use several demographics in our analyses, these are not the focal variables, and thus are not referenced in the results and discussion sections below. As noted in their discussion of ethical issues and dilemmas in learning analytics, Slade and Prinsloo (2013) share that “institutions should commit themselves to take due care to prevent bias and stereotyping, always acknowledging the incomplete and dynamic nature of individual identity and experiences” (p. 1524). The core premise in learning analytics then is to focus on those things which are actionable, particularly the institutional systems in which these courses are situated; thus, our focus here is on co-enrolled courses where students are likely to experience academic difficulty and the problems it poses across all students in our sample, while still controlling for the variance for which these demographics serve as a proxy.

Analysis
In a standard ordinary least squares (OLS) regression, it is assumed that the error terms in the regression should all have the same variance (i.e. consistent error variance). If these errors are not consistent, the OLS model is likely to provide incorrect estimates because of heteroscedastic errors. In our analysis, we hypothesize that the nature with which each of our independent variables of interest will impact the dependent variable (course grades by week) differs for each quantile into which a student’s grade would fall. In other words, experiencing below average academic performance in a co-enrolled course would impact a student with a grade in the 90th percentile in the focal course differently than a student with a grade in the 10th percentile. For the student in the 90th percentile, they may be able to recover from below average academic performance in co-enrolled courses without much harm to their grade in the focal course, while the student in the 10th percentile would likely experience additional pressures to recover in both the focal and co-enrolled courses.

Given the unequal relationships between dependent and independent variables across quantiles, we estimate our models using a quantile regression approach (Koenker, 2004). Specifically, we provide estimates for the relationships between dependent and independent variables across students in the 10th, 50th and 90th grading percentiles in the focal course. We used the “sqreg” package in Stata 15.1 for these analyses, which provides coefficients for the relationship between these variables for each quantile. For this method, the robust variance-covariance matrix of the estimators (VCE) was obtained via bootstrapping (100 iterations), allowing us to create confidence intervals comparing the different quantiles.

Note that the percentiles are not the grades themselves; instead, these estimates are normed around the average student performance in the course and should not be impacted by instructor conditions like grading curves or “difficult graders.” For example, a student in the 90th percentile would represent that student’s percentile rank when compared to their peers in the course, not a student that had a 90 percent grade in the course.

Results
The estimates from the quantile regression model for Biology and Economics, separated by term, are provided in Table II. Statistically significant results for each of the quantiles are italicized ($p < 0.05$), which provides a measurement of whether or not the coefficient is statistically different from zero (i.e. vs no relationship). The “sig” column to the right of the three quantiles in each model represents a Wald test of differences between each of the coefficients in the three quantile models.

Turning to the first independent variable of interest (i.e. disciplinary typology using Biglan’s, 1973 typologies), students in the 10th percentile with hard majors performed significantly poorer ($p < 0.05$) than their peers in soft majors for the Fall Biology ($\beta = -4.12$) and Fall ($\beta = -6.41$) and Winter ($\beta = -3.55$) Economics courses, while these same students
### Table II. Coefficients for quantile regression

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Biology – Fall</th>
<th>Biology – Winter</th>
<th>Economics – Fall</th>
<th>Economics – Winter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Q10</td>
<td>Q50</td>
<td>Q90</td>
<td>Sig</td>
</tr>
<tr>
<td>Female (vs Male)</td>
<td>−1.631</td>
<td>0.924</td>
<td>0.739</td>
<td>***</td>
</tr>
<tr>
<td>Race (vs White)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>1.224</td>
<td>1.985</td>
<td>−0.497</td>
<td>*</td>
</tr>
<tr>
<td>Hispanic</td>
<td>−8.244</td>
<td>−3.690</td>
<td>−2.738</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>−2.888</td>
<td>0.880</td>
<td>0.380</td>
<td>***</td>
</tr>
<tr>
<td>Multi</td>
<td>−1.338</td>
<td>−0.678</td>
<td>1.521</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>−11.436</td>
<td>−8.874</td>
<td>−0.172</td>
<td>***</td>
</tr>
<tr>
<td>International (vs US citizen)</td>
<td>5.280</td>
<td>−2.301</td>
<td>−3.251</td>
<td></td>
</tr>
<tr>
<td>First-year student</td>
<td>−5.502</td>
<td>−5.122</td>
<td>−1.056</td>
<td>***</td>
</tr>
<tr>
<td>Math placement test</td>
<td>1.122</td>
<td>0.750</td>
<td>0.359</td>
<td>***</td>
</tr>
<tr>
<td>Time indicators</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time2</td>
<td>−1.717</td>
<td>0.631</td>
<td>0.168</td>
<td>***</td>
</tr>
<tr>
<td>Time3</td>
<td>0.059</td>
<td>−0.018</td>
<td>−0.003</td>
<td>***</td>
</tr>
<tr>
<td>Curricular indicators</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biglan’s typology for majors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hard (vs soft)</td>
<td>−5.011</td>
<td>−4.117</td>
<td>−0.815</td>
<td>***</td>
</tr>
<tr>
<td>Pure (vs applied)</td>
<td>−1.885</td>
<td>−0.850</td>
<td>0.980</td>
<td>*</td>
</tr>
<tr>
<td>Life (vs non–life)</td>
<td>2.396</td>
<td>2.350</td>
<td>1.896</td>
<td></td>
</tr>
<tr>
<td>Undeclared major</td>
<td>−5.884</td>
<td>−3.429</td>
<td>0.123</td>
<td>***</td>
</tr>
<tr>
<td>No. of difficult courses</td>
<td>4.233</td>
<td>1.033</td>
<td>−1.424</td>
<td>***</td>
</tr>
<tr>
<td>No. of total co-enrolled courses</td>
<td>0.924</td>
<td>0.098</td>
<td>0.000</td>
<td>**</td>
</tr>
<tr>
<td>Hardships in co-enrolled courses</td>
<td>−5.378</td>
<td>−4.615</td>
<td>−2.154</td>
<td>***</td>
</tr>
</tbody>
</table>

**Notes:** $n = 1,785$. Statistically significant results for each of the quantiles are italicized ($p < 0.05$); Asterisks for Wald test of differences: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. 
in the Winter Biology course performed significantly better ($\beta = 3.48$). Looking across the quantiles, the magnitude of this relationship appears to slowly decrease to non-significant differences for students in the 90th percentile, except for in the Fall Economics course. The findings were less meaningful for the pure/applied designations, with only 10th percentile pure major students in the Fall and Winter Economics courses showing significantly lower scores than their peers in applied majors. Note, however, that there were significant differences between the coefficients across percentiles in three of the four courses ($p < 0.05$; except the Winter Economics course) suggesting differences in magnitude across these groups. Finally, for life/non-life designations, life majors in the 10th percentile did significantly better than peers who were non-life majors in the Biology courses, but performed worse than non-life majors in the Fall Economics course. The results for undeclared majors were mixed, with these students performing significantly worse in Fall terms (Biology = 10th and 50th percentiles, and Economics = 10th, 50th and 90th percentiles) and better for students in the 10th percentile of the Winter Biology course.

Students' co-enrollment in other difficult courses was only a significant factor for the Fall Biology and Winter Economics courses. For Biology, with every increase in one difficult course, students in the 10th ($\beta = 4.23$) and 50th ($\beta = 1.03$) percentiles actually did significantly better, while 90th percentile students did worse ($\beta = -1.42$). For Economics, students in both the 50th ($\beta = -2.61$) and 90th ($\beta = -1.03$) percentile did significantly worse with each difficult course. There were only significant performance differences across percentiles in the Fall Biology course ($p < 0.001$).

Next, the increased number of enrolled courses tended to actually benefit students in each of the percentiles given different courses and terms. The most consistent finding for the Fall and Winter Biology and the Winter Economics courses was that students in the 10th percentile were associated with a 0.82–1.06 percentage point increase in grades for each additional course enrolled.

Finally, the results were much stronger for students who experienced below average academic performance in co-enrolled courses. In nearly every percentile across all four courses (except for 90th percentile in the Fall Economics course), students experiencing below average academic performance in at least one co-enrolled course (regardless of course difficulty) did significantly worse in the focal course. For the 10th percentile students in the Winter Biology course, each additional co-enrolled course where a student showed below average academic performance was associated with a 13.28 percentage point drop in grades for the course ($p < 0.05$). Similarly, the same experience was associated with a 9.41 percentage point drop in grades for the Winter Economics course. The coefficients by percentile also significantly differed across all four courses ($p < 0.001$). For example, in the Winter Biology course, students in the 90th percentile who experienced below average academic performance in a co-enrolled course were only associated with a 0.67 percentage point drop in grades ($-0.67$) compared to students in the 10th percentile at $-13.28$ points.

Although the time indicators (i.e. changes in performance measured on a week-by-week basis) were not a focal point of interest in this paper, it is worth noting that the linear, quadratic, and cubic functions of time were significant across all quantiles ($p < 0.05$) and the differences across these quantiles were significant ($p < 0.001$). The implications for these findings are discussed in the section that follows.

Discussion
Returning to our research questions, we found several co-enrollment patterns that had a significant relationship with student success in two gateway courses. First, when using Biglan's typology to look at the disciplinary classification of students’ majors (RQ1a), we discovered that some of these classifications did appear to have a significant association with course performance when the student’s major was in a different classification than
the gateway course. For example, in the Economics course – a soft, applied, non-life course – students in “soft” majors performed significantly better when compared to students in “hard” majors, and “life” majors performed significantly worse than “non-life” peers. The results were opposite for the Biology course – a hard, pure, life course – with the exception of “hard” majors in the Fall Biology course. These findings mirror prior research, which suggests that Biglan’s typology can provide insight into how major fields organize knowledge, and therefore may help us understand how students’ prior educational preparation and interests may impact their performance across an array of gateway courses in a general education curriculum (Simpson, 2017).

Second, when taking into account a student’s performance quantile in the focal course (RQ2), the impact of co-enrollment in difficult courses (RQ1b) was mixed. Students in the 90th percentile performed significantly worse on their course grades in both the Fall Biology and Winter Economics courses when they were enrolled in at least one difficult course. Students in the 50th percentile in the Fall Biology course did significantly better than their peers with no difficult courses, but worse in the Winter Economics course. Students in the 10th percentile of the Fall Biology course who were enrolled in at least one difficult course did significantly better than their peers. While our prior research focusing on a computer science gateway course suggested a negative relationship between co-enrollment in difficult courses and college student success (Brown et al., 2018; DeMonbrun et al., 2018), the findings here indicate that the relationship between enrolling in a difficult course and gateway course performance is more complex than originally anticipated, as co-enrolling in difficult courses appears to positively influence students with lower course grades and negatively influence high achievers in the focal course. This perhaps counter-intuitive result suggests that underperforming students’ motivation and help-seeking (Zimmerman, 1990) may be increased when faced with multiple hard courses, while students who consistently do well may be less likely to work to overcome the competing demands of multiple difficult courses.

Third, the number of co-enrolled courses (RQ1c) did not have a strong relationship with course performance. For those students it did impact (i.e. mostly those in the 10th percentile), the percentage point increase in grades was by one point or less. This is somewhat unexpected given the research demonstrating the relationship between the number of credits taken and student success (Attewell and Monaghan, 2016; Calcagno et al., 2007; Hodara and Rodriguez, 2013); however, this also signals that the number of courses might not matter as long as students are performing well in all of them. It may also be that the students in our study, who have gained admission to a highly selective research university, provide a sample that is not representative of students in prior studies. For example, the freshman retention rate for the university proving data for this paper are significantly higher than the national average of 72 percent, and the students’ median household income is almost triple the national median of $59,039.

Fourth, the results suggest that experiencing below average academic performance in co-enrolled courses (RQ1d) was consistent across percentiles, as students demonstrated significant drops in focal course performance when experiencing below average academic performance in another (non-focal) course. Similar to the findings for difficult courses, the magnitude of these relationships was dependent upon the student’s performance in the focal course (RQ2), with a greater impact on students in the 10th percentile vs the 90th percentile. For example, the most dramatic association was for students who were in the 10th percentile of the Winter Biology course and experiencing below average academic performance in another course, as these students were predicted to score over 13 points lower than their same peers in the 10th percentile who were not experiencing below average academic performance in any other co-enrolled courses. This was in contrast to students in the 90th percentile of the same course, who were predicted to score 0.67 points below their peers who were not experiencing below average academic performance in co-enrolled courses.
This finding is consistent with our prior research in a gateway engineering course, which suggests that once students experience academic difficulty in one course, their grades in other courses are also susceptible to declines, especially for those students who are struggling relative to their peers in the same course (DeMonbrun and Brown, 2017; DeMonbrun et al., 2018).

Finally, as shown in Figure 1, there were significant differences in the pattern of student performance over the semester in the two introductory courses. There were significant drops in average course grades in the early weeks of the Fall semester for both Biology and Economics courses – week 3 for Economics and week 4 for Biology. Furthermore, students in the Fall Economics courses tended to recover, while Fall Biology students average course grades hovered around 81–83 percent for the remainder of the term. In the Winter term, student performance in both courses did not drop until week 7 and 8, respectively. This may reflect differences enrollment patterns, nuances in each course’s grading patterns or other exogenous factors (e.g. when the course was offered, differences in instructors) and demonstrate why it is important to conduct these analyses by course and term.

**Limitations**

There are several limitations to this analysis. First, this study was an examination of two courses over one academic year at one highly selective public institution in the American Midwest. Thus, these findings are more likely to be representative of similar types of schools (i.e. very-high research and selectivity) with similar student demographics. Additionally, although most of the introductory courses at this institution have large class enrollments, we did not consider how class sizes may impact student achievement. Second, we did not consider the impact that students’ motivation or their capacity for self-regulated learning would have on their performance in the current term. For example, perhaps a student experienced academic difficulty in a Biology or Economics course prior to this one (in either college or high school), and thus is motivated to perform better or accept doing worse in this course. The reverse could also be true with the student who has typically performed well in these courses in the past, but is struggling with the college-level difficulty of the course. These types of factors are beyond the scope of this study, but should be addressed in future work.

**Implications and conclusions**

Our prior work has shown that learning analytics can reveal how taking difficult courses simultaneously complicates students’ potential for academic success. This current work adds further evidence, but also complicates that story. The data presented here suggests that the decisions about which courses students should take when can significantly impact
academic achievement in introductory courses across various disciplines. The results point to factors related to the student (e.g. the alignment of their intended major to the introductory course) and the nature of their co-enrolled courses (e.g. how many are difficult for themselves and/or other classmates).

Our research offers implications for academic advisors and for academic planning in undergraduate higher education. First, when advisors consider the array of courses that students enroll in in the first year, they should encourage students to carefully consider the balance of skills and knowledge required to be successful in first year courses. For example, students who take introductory courses earlier in the curriculum may have a harder time responding to signals from performance assessments about their ability to be successful in the course as they are still potentially identifying resources on campus and developing strategies for completing academic work (broadly) and responding to academic difficulty (in specific courses). This is in contrast to students who are further along in their coursework, or who take an introductory course later than peers in the same academic year, as they appear more capable of managing difficult coursework. In this case, the most efficient path may not be the path to success, which in the event of deterred momentum makes that pathway no longer efficient. We know anecdotally that advisors already do this work with their students, but we would encourage institutions to support advisors by replicating this analysis with historical record data to help advisors supplement their knowledge of difficult introductory “weeder” courses with data-enhanced models showing patterns of co-enrollment that students should be encouraged (and discouraged) from pursuing. The goal here is not to prevent students from taking difficult courses. Rather, such an approach would encourage students to organize their courses to promote academic success.

Second, as a complement, we believe instructors, academic planners, and department leaders should also consider the impact of course scheduling for student success. Often, because of when courses are offered and the pre-requisite structure of major programs, students are left with few options but to take difficult courses simultaneously. If departments better understood what their students are doing concurrent to their enrollment in a difficult course in the department, they might be encouraged to make changes to when or how their courses are offered. Curriculum designers may even find themselves motivated to re-think curricular pre-requisites to support pathways that encourage student success and mitigate some of the difficulty students’ experience as a byproduct of their course enrollments.

College courses should be challenging to students, and it would be impossible (and potentially counter-productive to self-regulated learners) to develop student pathways toward degree that are frictionless. However, institutions can take a more proactive approach to understanding how students navigate the curriculum and what impact students’ course-taking decisions may have on their academic success. Although the results generated here may be related to this specific institution and the students who attend, the model we offer above provides a concise data-driven approach to translating theory about the organization of the curriculum into models for academic advising and curriculum planning in undergraduate higher education. This paper also provides a generalizable methodology for conducting learning analytics investigations that can be used by other universities to replicate this work and design curricular pathways for students at their institutions that have the potential to improve students’ success.

References


Hodara, M. and Rodriguez, O. (2013), Tracking Student Progression through the Core Curriculum, CCRC Analytics, Community College Research Center, Teachers College, Columbia University, New York, NY.


**About the authors**

Dr Robert Matthew DeMonbrun is Statistician for the Enrollment Management Research Group under the Office of the Provost at Southern Methodist University. Dr DeMonbrun’s program of research centers on teaching and learning in STEM education, particularly regarding the use of evidence-based instructional practices in the engineering classroom, as well as how student data collected from institutional sources can be used to understand student learning and performance in the classroom. Dr Robert Matthew DeMonbrun is the corresponding author and can be contacted at: mdemonbr@umich.edu

Dr Michael Brown is Assistant Professor in Higher Education and Student Affairs at the School of Education in the College of Human Sciences at Iowa State University, where he coordinates the Applied Research in the Human Sciences certificate program. His research focuses on the role of digital technology in academic planning, curriculum making and students’ academic success in undergraduate higher education. His work has been featured in *Computers and Education, The Internet and Higher Education, Review of Higher Education* and the *Journal of Learning Analytics*.

Dr Stephanie D. Teasley is Research Professor in the School of Information, Director of the Learning Education & Design Lab (LED Lab) and Core Faculty Member of the Michigan Institute for Data Science (MIDAS) at the University of Michigan. Her recent work has focused on assembling and utilizing institutionally held student data to design and evaluate new ways to support student success in Higher Education. From 2016 to 2018 she was the President of the Society for Learning Analytics Research (SoLAR), and she is currently the chair of the International Alliance for the Advancement of Learning in the Digital Era (IAALDE).

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Abstract

Purpose – The purpose of this paper is to answer the following questions: On which early alert system suggestions are students more likely to act? What factors drive students’ decisions to act on early alert system recommendations?

Design/methodology/approach – This study examined whether students’ behaviour changed after receiving the results of an early alert system (CDR). In the middle of a semester, 423 students with varying levels of English proficiency were invited to try the CDR and complete a questionnaire that asked about their perception of the tool and whether they planned to act on the recommendations they received.

Findings – Results suggested that students mainly planned to take the assessment-related recommendations provided through the CDR to improve their assessment performance. Results also suggested that student anxiety and student ability affected the likelihood that students would act on the recommendations.

Practical implications – These findings provide useful insights for early alert system designers to establish a system that generates useful recommendations for students.

Originality/value – The findings of this study contribute to the development of early alert systems. Designers can now realise what suggestions can be effectively offered to students.

Keywords Data mining, Learning analytics, Academic writing, Change of behaviours, Early alert systems, University teaching

Paper type Research paper

The use of early alert systems in education is common in this data-driven era, and falls under the spectrum of learning analytics. These systems were developed to identify at risk students and to give suggestions for students to improve. An overwhelming number of studies have reported that students improved after using these systems, such as Lawson et al. (2016) and Essa and Ayad (2012). However, previous studies have neglected to examine the process by which students respond to these systems and make the decision to improve (Howell et al., 2018; Wise, 2014).

Although learning analytic techniques attempt to use a data-driven approach to unveil the learning process (in addition to improving learning outcomes), previous studies focussed on the outcome of the suggestions rather than on the process of accepting the suggestions. This missing piece is important because practitioners need to understand why students accept the recommendations and what actions they take to improve after receiving an alert – for example, whether students are motivated to accept the system’s suggestions and what types of suggestions are most effective. Examining these factors may result in more effective suggestions being offered and may help to identify the psychological factors that facilitate students’ reactions to different types of suggestions.

Literature review

Learning analytics is not new in higher education. The most commonly used definition is from the First International Conference on Learning Analytics and Knowledge (2011): “the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs”.

The current issue and full text archive of this journal is available on Emerald Insight at: www.emeraldinsight.com/2050-7003.htm
Importantly, this definition indicates that learning analytics is different from traditional educational research because it focusses on how data-driven research informs teaching and learning processes, in addition to answering a research question.

The need to inform teaching gives rise to a number of learning analytics objectives (see Baepler and Murdoch, 2010, p. 5 for a full list). Predicting student outcomes is a learning analytics objective that focusses on predicting which students will be successful (Park et al., 2016) and which will be at risk (Klüsener and Portenbacher, 2015). Often, researchers attempt to predict students’ final grades using course variables, such as discussion forum posts (Romero et al., 2013) or other system indicators in the learning management system (Ruiz-de-Azcárate et al., 2017; Saqr et al., 2017). See Conijn et al. (2016) for an extensive list of potential prediction variables in the learning management system. Predicting student outcomes is an important task for post-secondary educators because classes are normally large, leaving little time to assess an individual student’s progress or provide feedback to individual students (Chingos and Whitehurst, 2012). However, if teachers know that particular students are struggling, they can provide timely help. Some teachers call their method of assisting struggling students “early warning systems” (Essa and Ayad, 2012) because they aim to alert students to problems as early as possible. With the problems identified, suggestions can be offered to help students improve.

Studies have adopted different approaches regarding suggestions offered by early alert systems. Lu et al. (2017) sent notifications to students or reminded teachers to arrange a face-to-face discussion with students, and concluded that these systems led to improved student engagement and outcomes. However, students’ perceptions of and reactions to the suggestions have not been reported. Chaudy and Connolly (2018) explored the usefulness of providing various indicators perceived by educators in the early alert systems; showing the amount of time spent by the student and participants’ progress reports were considered to be important.

Schumacher and Ifenthaler (2018) conducted one of the few studies to investigate students’ perception of learning analytics tools. They found that self-assessment feedback, learning recommendations and current status timeline were important features that students need. However, the study used an online survey and respondents did not actually experience an early alert system. This limits the generalisability of the study. Lonn et al. (2015) related students’ perceptions of an early alert system to the students’ own motivational aspects. They found that student perception of the tool was a strong predictor of motivation to change, and that if students did well (i.e. mastered the course), they were less likely to seek help from other resources suggested by the system. They also found that personal factors (such as family encouragement) and predicted course grades played roles in students’ motivation, and concluded that understanding students’ motivation for using early alert system tools was important. Howell et al. (2018) found that grades, but not expected grades, had a strong relationship with affective response to feedback and academic resilience. Students with higher grades had more positive affect and academic resilience, but were less likely to reflect or seek help than those who merely received pass or fail grades.

Research questions
The current study aimed to fill the gap in knowledge surrounding students’ perception of early alert system tools by exploring students’ behavioural changes after using the early alert system. The study addressed the following research questions:

*RQ1.* On which early alert system suggestions are students more likely to act?

*RQ2.* What factors drive students’ decisions to act on early alert system recommendations?
Methodology
Participants
Participants included 423 first- or second-year university students (210 male and 213 female) from all disciplines, including business, civil and environmental engineering and health sciences, who were enrolled in an English course.

All participants were high school graduates who had completed the secondary school exit exam, the Hong Kong Diploma of Secondary Education Examination, which assesses each subject using eight levels: unclassified, 1, 2, 3, 4, 5, 5* and 5**, and the results determine the university courses students can take. More than half of the study participants (57.4 per cent) achieved a Level 3 on the English exam, which is equivalent to the 5.48–5.68 range on the International English Language Testing System; the remaining participants achieved a Level 4 or above, equivalent to the 6.31–7.77 range or above on the International English Language Testing System (Hong Kong Examinations and Assessment Authority, 2012).

Participants were taking a three-credit English course with three assessments. The results of the first assessment had just been released, giving students their first assessment component grades. The course diagnostic tool was designed to use the first assessment component grades to make recommendations for the third assessment. More details on grade mapping and recommendations appear in a following section, “Course Diagnostic Report Tool”.

Procedures
The research team obtained permission from teachers to conduct the research in their classrooms. During the class visit, students were informed that their participation in the study was voluntary and the researcher introduced the Course Diagnostic Report Tool. Students were then given 5–10 min to try the tool, during which time the research team helped students who needed clarification. Students were encouraged to try the tool in at least two ways: entering their real grades to find how well they did in comparison with other students, and entering some fake failing grades to see specific recommendations for each assessment component. Next the researchers administered the questionnaire to all students, instructing those who did not wish to participate to return a blank questionnaire.

Course Diagnostic Report tool
The Course Diagnostic Report tool was designed to provide students with diagnostic feedback and encourage them to take action to improve performance on their upcoming assessments. The tool is basically a single web page, written in HTML using jQuery. The landing page offers an interface for students to enter their first assessment grades, including the overall assessment grade and the component grades, such as content, organisation, language and academic conventions (referencing) (see Figure 1 for a screenshot of the interface).

After entering the grades, students see their Course Diagnostic Report (Figure 2). The report includes grades entered, their performance compared with students in other cohorts and specific recommendations for improving their performance.

The first part of the tool, the results comparison, was inspired by Lonn et al. (2015) who suggested that course mastery affects whether students act on the suggestions and the comparison results indicate, to a certain extent, the level of course mastery. The comparisons generate one of three outcomes: above average, below average or bottom 5 per cent. If a student’s result is above average, a generic recommendation to visit the language centre website appears. If the student’s result is below average or in the bottom 5 per cent, the student receives specific recommendations for that component.

The second part of the tool includes learning recommendations suggested by Schumacher and Ifenthaler (2018) that may include online activities designed for the course (and embedded as a required element for the course), online resources offered by the
university language centre, books (with page numbers) related to the course and/or online courses designed by the language centre (e.g. the Learning4Life course on university study skills: https://edx.keep.edu.hk/courses/course-v1:PolyU+L4L+2018/about). Students may enter their e-mail addresses and forward the report to their e-mails for future reference.

The CDR’s development, wording and recommendations were confirmed with the subject leaders. For example, the subject leaders thought that using “average” to identify at risk students was more appropriate than using data mining techniques. Therefore, the whole site should meet the needs of students.

**Instrument**
A questionnaire was distributed during the class visit for collecting students’ comments on the tool and their information for further analysis (Appendix). The questionnaire consisted of eight parts: perception of the tool, tool-based action to be taken, additional action to be taken, attitudes about the course, linguistic self-confidence, classroom anxiety, demographic information and grades for the course and the previous exam (see Table I for details).

---

Figure 1.
Landing page for the Course Diagnostic Report

Figure 2.
The report diagnosis and recommendations
Data analysis

The questionnaire data were entered on a data sheet for further processing and analysis. Any questionnaires with one whole section left blank were excluded. To analyse the data more efficiently, a mean score was computed for each aspect (i.e. aspects 1–6 in Table I) to better understand students’ attitudes regarding those aspects.

The first part of the analysis was conducted with IBM SPSS Statistics, version 25 to examine students’ comments on the tool itself. Then, a decision tree analysis was conducted using R Studio (Version 1.1.456; R: version 3.4.3) to identify students’ determination factors following the system recommendations. Generally, a decision tree can be used to solve a classification problem (Tan et al., 2014) by identifying factors that predict the outcome, i.e., whether students will accept the suggestion. Decision tree is a popular technique for prediction because of its simplicity and comprehensibility for different types of data structures (Shahiri et al., 2015). Romero et al. (2008) also revealed that decision tree is simple reasoning process and the convenience in generating if ten rules and these are important to the current study.

All measures were converted to categorical variables for analysis. The cut-off scores were based on the statements the participants rated. Details of these statements are shown in Table II.

Two rounds of analysis were conducted to predict whether students accepted tool-based suggestions, and general recommendations. A standard “holdout method” was used for ensuring the validity of the results (Tan et al., 2014). The original data set (with all records) was divided into a training data set (282 records) used to develop the decision tree, and a testing data set (141 records) used to validate the results. The two decision trees were developed separately, i.e., training/test data sets were prepared for one decision tree, and
different training/test data sets were prepared for the other decision tree. The error rate of the testing set was used for analysis (i.e. the number of wrongly predicted records divided by the total records in the data set).

Results

Overall perception of the tool

The first section of Table III shows the descriptive statistics. Among the five elements being assessed, the students positively perceived all of them, with all items having scores above 3 (on a five-point Likert scale). The students found the tool useful and would like to have a
similar tool in other courses. Of the factors considered, the user-friendliness rating was the highest (mean = 3.87). The usefulness of the suggestions had the lowest rating among the scales with a mean of 3.29, which was lower than the rating on the usefulness of the results comparisons (mean = 3.49).

Types of suggestions likely to be accepted
The first research question focussed on whether students would accept any of the suggestions. The mean scores for online course-based suggestions were 3.11 (for course-based online quizzes) and 3.28 (for course-related online resources). Suggestions to read books and enrol in online courses were rated lower, with means of 2.91 and 3.08, respectively. Hence, it was less likely that students would accept the latter two suggestions.

When students were asked to indicate whether they were likely to work harder in the course in other aspects, their ratings were lower. Among the four suggestions, only the mean score for “more time for preparation of next assessment” reached 3 (i.e. likely). Other suggestions, such as “being more attentive”, “meeting with the teacher” and “doing more exercises to improve proficiency”, scored lower, indicating decreased likelihood of students accepting these suggestions.

Factors driving improvement
Figure 3 shows the results of the decision tree analysis. Each box is called a node and the first line indicates the node category (i.e. likely or unlikely to accept the tool-based recommendation). In the second line, the number on the left indicates the number of records that fell into the node’s category (i.e. likely or unlikely, as indicated on the first line). The last line indicates the percentage of records, among the total number of participants in this data set, that fell into this node. For example, as indicated in the first node on the right at the bottom, 238 of 282 participants in the training data set were likely to accept the tool-based recommendation and likely to accept the general recommendation, whereas 15 participants were not likely to accept the tool-based recommendation but were likely to accept the

Figure 3. Decision tree to predict the likelihood of taking tool-based recommendations
The total records in this node (238+15 = 253), accounted for 89.7 per cent of the total records. The error rates for this tree are presented in Tables IV and V.

The error rate of the testing data set (13.5 per cent) was acceptable and comparable to the ten data mining studies on predicting student’s outcomes summarised in Shahiri et al. (2015, p. 418), with a mean of 21.1 per cent ranging from 9 to 35 per cent. For tool-based suggestions, the most important predictor was accepting the general recommendation. Participants, who accepted the general recommendations (90 per cent of all participants), were likely to accept the tool-based suggestions as well. Of those who did not accept the general recommendations, only those who had a positive attitude towards the tool and/or average English proficiency were likely to accept the tool-based suggestion.

The decision tree for general recommendations is presented in Figure 4.

<table>
<thead>
<tr>
<th>Table IV.</th>
<th>Predicted Likely</th>
<th>Actual Likely</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likely</td>
<td>252</td>
<td>18</td>
<td>270</td>
</tr>
<tr>
<td>Unlikely</td>
<td>1</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>253</td>
<td>29</td>
<td>282</td>
</tr>
<tr>
<td>Notes:</td>
<td>Accurate prediction = 252+11 = 263; wrong prediction = 1+18 = 19; error rate = 19/282 = 6.7 per cent</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table V.</th>
<th>Predicted Likely</th>
<th>Actual Likely</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likely</td>
<td>120</td>
<td>18</td>
<td>138</td>
</tr>
<tr>
<td>Unlikely</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>121</td>
<td>20</td>
<td>141</td>
</tr>
<tr>
<td>Notes:</td>
<td>Accurate prediction = 120+2 = 122; wrong prediction = 1+18 = 19; error rate = 19/141 = 13.5 per cent</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Figure 4.** Decision tree for accepting general recommendations
General recommendations were most likely to be accepted by those who accepted the tool-based recommendation. Of those who did not accept the tool-based suggestions, having a positive attitude to the tool and the course made them more likely to accept the general recommendation. The error rate of the testing set (11.3 per cent) was acceptable as well, comparing the results summarised in Shahiri et al. (2015) (see Tables VI and VII for more details).

Discussion

The study examined how students perceive the recommendations from an early alert system and what types of recommendations were more likely to be acted on. In addition, factors driving the decisions regarding whether to act on recommendations were examined.

The results indicate that tool-based recommendations are more likely to be accepted by students and that course-based online resources and online activities are the most popular recommendations. These two suggestions are presented as ways to improve their assessment performance, which may be a reason that students are more likely to act on these recommendations. In fact, a general recommendation to spend more time on assessments was also likely to be acted on, perhaps reflecting an assessment-driven mindset, which may affect their decisions regarding early alert system recommendations. This echoes what Schumacher and Ifenthaler (2018) reported as learning recommendations to be wanted by students. This also aligns with what Schumacher and Ifenthaler (2018) revealed as students’ need for skills enhancement feedback. Interestingly, course-related books were recommended but this recommendation was not welcomed by students. Perhaps students are very “pragmatic” in doing extra work for their courses (Huon et al., 2007) and do only tasks that can effectively improve their assessment grades. Reading course-related books may not be perceived as a useful and effective way to improve, resulting in students being less likely to act on these recommendations.

The implication of this is obvious and profound. Suggestions for analytics system should be assessment-related and effective enough to help students do well in the forthcoming assessment. Despite the usefulness of other types of recommendations (e.g. course readings), they may not be effective in an analytics system because students appear reluctant to act on them. To develop effective systems, early alert system designers need to include recommendations that students are likely to act on (Gašević et al., 2015).

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Actual</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likely</td>
<td>252</td>
<td>13</td>
</tr>
<tr>
<td>Unlikely</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td>Total</td>
<td>256</td>
<td>26</td>
</tr>
</tbody>
</table>

Notes: Accurate prediction = 252+13 = 265; wrong prediction = 4+13 = 17; error rate = 17/282 = 6.0 per cent

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Actual</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likely</td>
<td>121</td>
<td>11</td>
</tr>
<tr>
<td>Unlikely</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>126</td>
<td>15</td>
</tr>
</tbody>
</table>

Notes: Accurate prediction = 121+4 = 125; wrong prediction = 5+11 = 16; error rate = 16/141 = 11.3 per cent
Regarding the factors that potentially drive students’ decisions to accept system recommendations, the current study showed that acceptance of recommendations in one area affects acceptance in another area. Other factors also affect recommendation acceptance, including, attitudes towards the course and ability (i.e. English proficiency in this context). It is logical that ability is an important predictor for students’ decisions surrounding recommendations for improving. If students have an average proficiency, they see their potential and are motivated to improve. These findings are consistent with Lonn et al. (2015), who found that students will seek help if they see themselves unable to master the course skills. Course attitude is also a logical predictor of accepting the general recommendations, which involve many course-related elements, such as being attentive in class and talking to the teacher. If the student has a positive attitude in class, he or she is likely to accept these recommendations.

Pitt and Norton (2017) concluded that expected grade creates a stronger response than the absolute grade. However, in the current study, expected grade was not identified as a predictor in the decision tree analysis, which is similar to the results reported by Howell et al. (2018). Students’ expectations in the current study did not vary a lot; 93.6 per cent of students expected their grade to be between a C+ and a B+, which may have affected their response to recommendations. Having said that, the author believes that this represents the population because the majority of students fall within the above grade range.

The overall implication is that recommendations provided by early alert systems need to convince students that it is possible to improve regardless of their existing ability. Similar to other studies that advocate the needs to pay attention the impact of academic resilience and disheartening negative feedback (Roberts et al., 2016), the current results suggest that students will be demotivated if they see no hope. If they are reminded of their existing competence (e.g. English exam scores in this study), it is more likely for them to have hope and act on the offered recommendations.

Conclusion
This study explored behavioural changes after the use of an early alert system known as the CDR. Similar to other educational interventions, after receiving the CDR, students tended to explore ways to quickly improve their assessment component grades. Other measures (such as reading a book) that were not directly related to the assessments were not popular and were less likely to induce behaviour changes. It was also notable that students who had good expectations for their grades and who perceived the report positively were likely to follow the provided suggestions.

Limitations
This study has several limitations; therefore, readers should interpret the findings with caution. For example, the sample size was relatively small, which weakens the generalisability of the results. However, with more than 30 samples, the study findings should still be trustworthy. In addition, similar to most analytics studies, results can differ across different contexts. For example, the findings on the assessment-driven mindset of students may be more relevant to Asian learners but not to other learners. The results are still a meaningful starting point for context and comparison with future studies involving diverse populations.

Contributions
The findings of this study contribute to the development of early alert systems. Designers can now realise the student perception towards different types of suggestions and develop useful suggestions that students are likely to accept.
References


Tan, P.N., Steinbach, M. and Kumar, V. (2014), Introduction to Data Mining, 1st ed., Pearson Addison Wesley, Boston, MA.

# Appendix. Questionnaire

## Language Enhancement Grant Project

**Course Diagnostic Reports – Evaluation Questionnaire (ELC 1012 / 1014)**

### Part A – Comments on the Tool

<table>
<thead>
<tr>
<th>Items</th>
<th>Strongly Disagree</th>
<th></th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The interface of the tool is user-friendly</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>2. The results of the comparison were useful</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>(e.g. “Above Average”, “Below Average” and “Bottom 5%”)</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>3. The suggestions provided were useful</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>4. The tool allowed me to know my strengths and weaknesses</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>5. I wish I could have a similar tool for other courses</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

**After using the tool, I will...***(Not applicable = Not on my list of recommendations)***

<table>
<thead>
<tr>
<th>Items</th>
<th>Highly Unlikely</th>
<th>Unlikely</th>
<th>Likely</th>
<th>Highly Likely</th>
<th>*Not Applicable</th>
</tr>
</thead>
<tbody>
<tr>
<td>6. do the IndiWork activities suggested by the tool</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>7. read the books recommended by the tool</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>8. use the online resources suggested by the tool</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>9. enroll in the online courses recommended by the tool</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

**After using the tool, I will...**

<table>
<thead>
<tr>
<th>Items</th>
<th>Highly Unlikely</th>
<th>Unlikely</th>
<th>Likely</th>
<th>Highly Likely</th>
</tr>
</thead>
<tbody>
<tr>
<td>10. spend more time to prepare for the next assessment</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>11. be more attentive during class</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>12. make an appointment with my English teacher to talk about my progress</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>13. do more exercises to improve my general English proficiency</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>14. What other functions/features do you think this tool should have?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Any other comments?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Part B – My Motivational State

<table>
<thead>
<tr>
<th>Items</th>
<th>Never</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I wish we had more English lessons this semester</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>2. I like English lessons this semester</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>3. English is one of my favorite subjects this semester</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>4. When the English lesson ends, I often wish it could continue</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>5. I want to work hard in English lessons to make my teacher happy</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Items</th>
<th>Never</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>6. I enjoy my English lessons this semester because what we do is neither too hard nor too easy</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>7. I would rather spend time on subjects other than English</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>8. Learning English is a burden for me this semester</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>9. In English lessons this semester, we are learning things that will be useful in the future</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>10. I feel I am making progress in English this semester</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>11. I believe I will receive good grades in English this semester</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>12. I often experience a feeling of success in my English lessons this semester</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>13. In English lessons this semester, I usually understand what to do and how to do it</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>14. This semester, I think I am good at learning English</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>15. I am worried about my ability to do well in English this semester</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>16. I often volunteer to speak in English lessons</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>17. I get very worried if I make mistakes during</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>
Shaping minds and changing behaviours

English lessons this semester
18. I am afraid that my classmates will laugh at me when I have to speak in English lessons □ □ □ □ □
19. I feel more nervous in English class this semester than in my other classes □ □ □ □ □

Part C – My Information (*Circle the appropriate option)

Gender  * Male/ Female

Programme of study

My expected overall grade for this course

My HKDSE English Results:
(Try your best to recall all your grades)

Reading
Writing
Listening and Integrated Skills
Speaking

My Other English Exam Results:
(please fill in this part if HKDSE is not available)

Name of the Exam:
Writing:
Speaking:
Overall

Part D – Informed Consent

Other than the information provided above, I sign below to give my consent to the research team to obtain my final grade and to conduct further research for teaching and learning purposes. All information provided will remain anonymous.

My Full Name
My Signature

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Abstract

Purpose – An ideal learning analytics tool for programming exercises performs the role of a lecturer who monitors the code development, provides customized support and identifies students at risk to drop out. But a reliable prediction and prevention of drop-out is difficult, due to the huge problem space in programming tasks and variety of solutions and programming strategies. The purpose of this paper is to tackle this problem by, first, identifying activity patterns that indicate students at risk; and, second, finding reasons behind specific activity pattern, for identification of instructional interventions that prevent drop-out.

Design/methodology/approach – The authors combine two investigation strategies: first, learning analytic techniques (decision trees) are applied on features gathered from students, while completing programming exercises, in order to classify predictors for drop-outs. Second, the authors determine cognitive, motivational and demographic learner characteristics based on a questionnaire. Finally, both parts are related with a correlation analysis.

Findings – It was possible to identify generic variables that could predict early and later drop-outs. For students who drop out early, the most relevant variable is the delay time between availability of the assignment and the first login. The correlation analysis indicates a relation with prior programming experience in years and job occupation per week. For students who drop out later in the course, the number of errors within the first assignment is the most relevant predictor, which correlates with prior programming skills.

Originality/value – The findings indicate a relation between activity patterns and learner characteristics. Based on the results, the authors deduce instructional interventions to support students and to prevent drop-outs.

Keywords Motivation, Drop-out, Learning analytics, Programming course

Paper type Research paper

Introduction

Drop-out in computer science is a great challenge. Drop-out rates of over 30 percent at course level (Watson and Li, 2014) and over 40 percent at program level (Heublein, 2014) are common. Reasons for drop-out in computer science are manifold. False expectations on content and topics, problems with performance requirements, as well as time restrictions based on student jobs are some of the most important reasons (Isleib and Heublein, 2017).

One of our expectations regarding learning analytic systems is that they support teachers by accompanying learners individually on their way and help to prevent drop-out. The learning analytic system has to cope with two questions in this context: Does a student need support in his/her learning process and what kind of instructional intervention could help? In the case of non-numerical or non-binary inputs – like texts, mathematical equations or program code – answering these questions is difficult. First, the representation of the student’s knowledge level is a multidimensional information set. Additionally, the related input sets must be evaluated as process features (duration of operation, frequency of activities, style, complexity, fragmentation, etc.). Ideally, a lecturer is aware of the most important features and good support strategies. According to different characteristics of
programming courses, their rules, expected and available skills, specific learning goals, etc., the relevant features are widely different.

Previous work related to the automated evaluation of student activities does not cover the complex problem en bloc. While several studies identified predictors for student’s performance from their programming behavior (e.g. Watson et al., 2013), the prediction of drop-out is much less researched (Falkner and Falkner, 2012). Furthermore, the question of appropriate instructional interventions based on these results is barely studied (Ihantola et al., 2015). For provisioning effective instructional interventions to prevent drop out, the underlying cognitive and motivational mechanisms have to be identified. To approach this challenge, learner characteristics, e.g., prior knowledge, or time management strategies, as well as demographic data have to be considered (Lau and Yuen, 2011).

Consequently, this paper deals with two questions:

1. Which variables from programming behavior can predict drop-out?
2. Can we find a relation between these variables and learner characteristics?

The contribution of this work is an analysis of indicators for drop-out and the deduction of appropriate countermeasures. Therefore, we discuss existing prediction strategies and feature sets, described in the literature, and presents results from a study on drop-out prediction and identification of related learner characteristics. Based on the findings, we discuss possible support strategies.

State-of-the-art

With the help of learning analytic approaches, we want to detect students at risk in a semi-automatic manner and preferably early in the course (Papamitsiou and Economides, 2014). Several studies analyze programming behavior for data-driven prediction of performance. Three different classes of features and metrics for assessing program code of students can be identified:

1. code style (e.g. number of lines, variable names, complexity (Edwards et al., 2017));
2. the appearance of errors during compilation, e.g., error types, error frequency, the sequence length of continuous errors; and
3. the programming/debugging process, e.g., the frequency and the number of lines changed per commit or activities after failed compilation (Altadmri and Brown, 2015; Munson and Schilling, 2016).

Especially algorithms on the basis of errors and error handling, like the Error Quotient (Tabanao et al., 2011) and the Watwin Score (Watson et al., 2013), can predict performance quite well. Jadud (2006) and Jadud and Dorn (2015) presented the Error Quotient concept that generates a scalar value representing the length and the structure of an error sequence. Watson et al. (2013) and Becker (2016) extended this approach. Öztürk et al. (2018) combined these approaches and determine benefits and drawbacks related to a specific programming course. However, also simple variables like total number of errors (Rodrigo et al., 2009; Tabanao et al., 2011), number of error streaks (Rodrigo et al., 2009) or time between compiles (Munson, and Schilling, 2016) are appropriate predictors of performance in a course. These studies focus on performance measures, usually the grade in the exam, with the goal to predict students that perform poorly or fail in exam. They do not consider students who drop out early in the course. Studies concerning data-driven prediction of drop-out during a course on the other hand so far often focus on analysis of LMS-data and similar systems and do not take into account programming behavior. An early start (Koprinska et al., 2015; Vihavainen et al., 2013) and timely submission of students’ assignments (Falkner and Falkner, 2012) as well as activities in social learning, like participation in discussion boards (Koprinska et al., 2015), are good predictors here.
To prevent or minimize drop-outs a possible approach is to implement appropriate instructional interventions which support students at risk. However, a literature review revealed that studies that employ their data-driven results to derive instructional interventions are scarcely to find (Ihantola et al., 2015). This is certainly due to the complexity of this goal. We want to explain this with the example of compiler errors (Öztürk et al., 2018). Errors are part of the learning process, especially in learning programming. However, the occurrence of too many errors or error streaks is an indicator for dropout-prone students (Öztürk et al., 2018). The reason behind too many errors can lie in cognitive states of the students like cognitive overload or misconceptions based on insufficient prior knowledge (Hawlitschek et al., 2019). Apart from that, effects on motivation are possible. If students are stuck in an error streak (i.e. getting the same error message repeatedly) and they cannot handle it, this could be a frustrating experience (Rodrigo and Baker, 2009), leading to drop-out. To be effective an instructional intervention has to be based on cognitive or motivational characteristics, states or demands of the user, which might be the causes for dropout-prone programming behavior. In this regard, we can distinguish between cognitive or motivational factors that are prior to and factors that occur or are affected during the programming process. In this paper, we focus on the former, which we refer to as learner characteristics. In literature on instructional design, different learner characteristics are examined, in particular cognitive variables like prior knowledge (Tseng et al., 2008) and cognitive or learning styles (Leutner and Plass, 1998; Plass et al., 1998). Cognitive theories, which focus on memory and information processing, like Cognitive Load Theory (Plass et al., 2010), consider domain-specific prior knowledge as especially relevant for the learning process. Depending on prior knowledge, the learner needs more or less support to process the learning content and to avoid cognitive overload or boredom (Kalyuga et al., 2001). Studies also reveal the effect of prior knowledge on drop-out in computer science courses (Horton and Craig, 2015). Competencies like bug-fixing and program understanding are closely related to prior knowledge and experiences (Wilson and Shrock, 2001). Studies reveal that students with low prior knowledge, which learned to program without any type of instructional guidance had problems to understand underlying concepts of programming (Mayer, 2004).

Furthermore, how persistently, how intensive and how often learners deal with learning content depends on their motivation (Rheinberg, 2008). This is also true for the decision to drop-out in computer science courses (Kori et al., 2015; Kori et al., 2016). The motivation to learn can be understood as the result of the interrelation between the learner and the situation. Self-efficacy expectations play a central role for the valuation of the situation and the engagement in the learning task (Salomon, 1983). The perceived self-efficacy depends on learners’ belief that they have the competencies to handle a learning task. Poor self-efficacy leads to poor engagement, satisfaction and performance (Brosnan, 1998; Liaw, 2008; Salomon, 1983) and is related with drop-out (Pintrich, 1999).

In our study, we analyze which motivational, cognitive or demographic variables correlate with relevant programming behavior. If we can identify a relation between programming behavior that is an indicator for drop-out and learner characteristics, the automatic provision of customized support becomes possible.

Study

Description of the course

This investigation is based on data obtained from an Embedded Systems course at a German University held in 2017/2018, where students had to accomplish five programming tasks, with increasing difficulty, by using the programming languages C++, C and assembly. The tasks (aggregation and evaluation of digital and analogous values, controlling of motors, and collision-free movement of small robots) build upon and complement each other. Hence, only students that successfully submitted every assignment, one task in two weeks, were allowed
to participate in the final exam. Tutors realized the evaluation of the task manually. From 67 undergraduate students, 38 passed the course successfully.

Courses were developed with LiaScript, an open-source authoring and development tool for interactive educational content (Dietrich, 2019). The open framework is currently hosted at https://LiaScript.github.io and can be used to render any kind of OER that is based on Markdown. The screenshot in Figure 1 depicts a similar and open course on CSharp[1] that was developed at the University in Freiburg.

Data mining technique
Each action of the students in the programming environment was logged and stored in a database. If a student starts or finishes the application, pushes “save,” “compile” or “execute” buttons or enters the documentation, the system records the following criteria: programming history, version changes, switching between changes (students could manually go back to previous versions), compile times and results, changing focus on program files, changes in rendering mode (minimized/mid-size/full screen) per file. Next to programming, the results for quizzes and surveys were also stored. At the end of the one semester course we received 2.2 GB of data.

Note, we already aggregated our times series data in a first data preprocessing step. Nevertheless, the data dimensionality is too high, to easily identify patterns. Therefore, we have to reduce the dimensions (features) by using data mining techniques. The time series of actions and code versions for each student converted in a set of characteristic values summarizing the individual learning process. This “fingerprint” covers coding style (number of lines changed between two versions), code quality (occurrence of errors, error counts), schedule parameters (starting time after task publication) and the final success/failure. The logged variables are listed in Table AI. The chosen set of features is motivated by previous work from (Altadmri and Brown, 2015; Öztürk et al., 2018; Koprinska et al., 2015; Tabanao et al., 2011; Watson et al., 2013; Vihavainen et al., 2013). At the end, this data set for programming behavior includes 23 features for 67 participants.
At a second stage, we applied classification trees from machine learning, due to the fact they are easy to compute, but also easy to interpret by humans. Furthermore, they can be used as qualitative predictors with mappings to a rule-based system. Nevertheless, it has to be mentioned that the predictive accuracy is lower compared to other methods. Using aggregation methods, like random forests or boosting, might overcome this issue. For further details, we refer to Witten et al. (2011) and Han and Kamber (2006).

Additional questionnaire

Next to logging questionnaires were used to examine factors that are prior to the programming process. Based on Park and Choi (2009), we distinguish between factors that are external to the learning process (e.g. time constraints, resulting from job occupation) and factors which are connected with learning (especially prior knowledge/skills of the learner). Therefore, in the first lecture, before the exercises, the prior knowledge was administered with a test, consisting of 22 multiple choice questions on different themes on the course and two questions were students had to estimate the results of given code. The absolute test score for each student was used for further analysis. Additionally, students had to indicate their programming experience in years and their perceived programming skills as well as their experience with C, C++, Java and Assembly (Siegmund et al., 2014). We also measured their motivational beliefs with a rating scale (Cronbach’s α: 0.85) on self-efficacy in programming (items based on Compeau and Higgins, 1995; Holden and Rada, 2011; Law et al., 2010). We used a Likert-type rating scale ranging from 1 (very low/strongly disagree) to 5 (very high/strongly agree). As prior external factors, we asked the students to estimate the approximate time they spent in the week for lessons on-campus, self-studying and student job. We present an excerpt and description of variables derived from our questionnaire in Table AII.

Extraction of relevant features for drop-out prediction

We applied several decision tree algorithms, e.g., C 4.5 (Quinlan (1993), CART (Breiman et al., 1984), AID (Sonquist and Morgan, 1964) and restrict our presentation to C 4.5, because the results did not differ significantly. We assume that students who leave the course at an early stage have a different motive than students that leave later. For our machine learning approach, we distinguish between students who drop out in the first exercise (early drop-outs) and students that drop out between second and fourth exercise (later drop-outs). We only use data available at time (\(t = 1\)) when the first exercise was finished, according to our goal to find early predictors.

Results

Predicting a drop-out right after (or while doing) the first exercise was observed on 10 students who dropped out and 57 students who also started exercise 2. We used the RWeka (Hornik et al., 2009) implementation J48 from R Core Team (R Core Team, 2014) and depict the resulting tree in Figure 2.

Three variables describe our drop-outs after the first exercise. The most important variable is the delay between publishing task and the first compilation within the system. Already four students, who begin later than eight days do not finish the first exercise. Another important aspect is the count of successful commits and the error ratio. Note, about 95 percent of instances are correctly classified for this decision tree.

For the analysis of students who drop out at a later stage, we also use our measurements from the first exercise. We exclude students that drop out before the second exercise from this analysis, Figure 3 depicts the resulting decision tree for obtained by the J48 algorithm. Note, the classification quality decreases to about 86 percent. For those students, the error count differentiates drop-outs quite good. Additionally, the maximal length of an error
Figure 2. Decision tree (J48) with drop-out at \( t = 1 \)

Figure 3. Decision tree (J48) with drop-out at \( t \) in \( (2, \ldots, 4) \)
streak (S_trans_max) is useful to distinguish between drop-outs and students who finish the fourth exercise.

In a second view, we want to analyze the relationship of the variables identified in the decision trees and the motivational, cognitive and external factors from our questionnaire. The aim of this analysis is to identify possible interventions, that reduce drop-outs in the future. Due to the limited number of observations, we restrict ourselves to a correlation analysis and present Pearson-correlation coefficients for variables identified in the decision trees.

In the first step, we consider variables, which predict drop-outs after the first exercise (all participants) from the derived decision tree. We analyzed correlations between these variables and the cognitive, motivational, as well as external variables that we measured prior to the first exercise. We found significant correlations between the delay of students starting the exercise (Tn_max_delay) and programming experience in years ($r = 0.39, p = 0.00$) as well as time per week for student job ($r = 0.35, p = 0.01$). Students with more programming experience and more working hours in their job tend to start the exercise later. For the number of successful compiles in the exercise (success count), we found significant correlations with perceived programming skills ($r = -0.29, p = 0.03$). Although not significant, the correlation with experience in Assembly ($r = 0.24, p = 0.08$) and in C ($r = 0.25, p = 0.07$) is worth mentioning, because it gives a clue about possible lacks in programming skills. The more experience in Assembly or C students had, the more successful compiles they had. Furthermore, we found significant correlations between error ratio and self-efficacy ($r = -0.27, p = 0.04$) as well as prior knowledge ($r = -0.30, p = 0.03$) and estimated time in lessons on campus ($r = -0.27, p = 0.05$). The lower the perceived self-efficacy, the prior knowledge and the time in lessons on campus the higher the error ratio. These findings can also be relevant for future instructional interventions.

Second, we analyzed correlations between the variables, which predicted drop-out in the second decision tree (without students that dropped the first exercise) and the cognitive, motivational and external variables. For the numbers of errors in the programming process (error count), we found significant correlations with perceived programming skills ($r = -0.34, p = 0.02$) and experience with Java ($r = -0.34, p = 0.02$). The lower the perceived programming skills and the experience in Java the higher the error count. The maximal length of an error streak (S_trans_max) is a predictor for drop-out in the second decision tree as well. We identify significant correlations between S_trans_max and estimated time in lessons on-campus ($r = -0.43, p = 0.00$), programming experience in years ($r = -0.30, p = 0.04$) and prior knowledge ($r = -0.37, p = 0.01$). The lower the programming experience as well as the prior knowledge and the less time the students on average spend in lessons the longer is the maximal duration of an error streak.

Discussion
The aim of our empirical study is to identify variables from programming behavior that can predict drop-out and to find relations between these variables and learner characteristics. In the following, we discuss our results and deduce instructional interventions. Students, who drop out during the first exercise, apparently have different reasons than students, who drop out later. The user behavior variables that indicate drop-out for both classes of students are different. The time between start of the exercise and the first-time students logged into the system to work at the task (Tn_max_delay) is the best predictor for students to drop out in the first exercise. The same findings are reported in literature (Koprinska et al., 2015; Vihavainen et al., 2013). The positive correlation between programming experience in years and delay time can be either explained with an overestimation on one’s own capabilities or it is a problem of student’s time management. We suppose that students with more programming experience approach the first exercise with an underestimation of required time they need to solve the task. We draw this explanation from the positive
correlation between time on student’s job and delay time. The correlation between programming experience in years and time for students’ job per week \( (r = 0.43, p = 0.00) \) also indicates a possible mixture of both explanations. Because of their prior experience, students seem to start later. If the exercise demands more resources than expected, scarce time resources due to a student job lead to drop-outs additionally. However, while we can only guess the true reasons for a delayed start of the exercise, the intervention is an easy one: we propose to send a short message to students with more than two days of delay, which points out the relevance of an early start. Students with jobs also seem to have problems meeting the time requirements. A practical approach could be to give more flexibility to students with occupational burdens.

Additionally, success count and error ratio are relevant variables to predict drop-out. A small number of success counts indicates that a student is at risk to drop out. However, success count as predicting variable for early drop-outs has to be handled with care. This result in the decision tree is biased, because it results from drop-out itself. Students that only tried a very short time to solve the exercise before they drop out naturally have a small number of success counts. On the other hand, for students who try to solve the exercise, a small number of successful compilations indicate problems with programming. To handle this problem, the experience with programming languages C or Assembly as well as prior knowledge should be considered. If the number of successful compilations in the first quarter of the processing time deviates from the median more than 1.5 sigma (Wheeler, 2004), the lecturer can provide guidance based upon students’ needs. Another way, to prevent early drop-out, is the preparation of preparatory programming course in Assembly or C. Based on a questionnaire to detect students with unsatisfactory programming skills a preparatory course could automatically be provided. In this way, necessary prior skills for solving the first exercise are obtained and at the same time, heterogeneity of students programming experience is reduced.

The second analysis is even more relevant for our goal to prevent drop-out in the course. These are students, who struggled through the first exercise and dropped out later. If we identify variables that give us clues on potential drop-outs already in the first exercise, we have the opportunity to help these students to successfully complete the course. In our decision tree (Figure 2), the error count is the most selective variable for predicting later drop-outs. This output replicates results from literature (Tabanao et al., 2011). Thus, the variable is a lead indicator for the lecturer. For students with an error count above the median, instructional interventions should take place. A first intervention idea here: the lecturer simply asks these students for problems with the exercise. Additionally, as a result from our correlation analysis, we know that the higher the perceived programming skills and the experience with Java, the lower is the error count in the first exercise. Hence, Java could also be a valuable part of an instructional intervention. The results points on the necessity to identify prerequisites in programming that all students need to complete the course successfully and to implement them as preparatory online-course. However, to prove the effectiveness of this approach requires further data collection and analytics.

The maximal length of an error streak is also a relevant predictor. The correlation analysis indicates that students who get stuck in an error streak have fewer programming experience and prior knowledge. Especially knowledge on compiler error messages seems to be the problem here. The students apparently are not able to correct an error after a first error message. A better description of error messages for those students or a more prominent marking of the relevant code line (Pedroni and Meyer, 2008) should be applied. This additional help should be provided only for students that are stuck in an error streak, to make the exercise not too easy for all other students. Furthermore, the estimated time students spent per week in lessons on-campus correlates negatively with the maximal length of an error streak. To clarify the benefit of lecture participation with reference to such results is a possible intervention here.
To sum up, we consider the implementation of learning analytics a fruitful method to prevent drop-out in programming courses. A prediction based on decision trees identifies variables that go hand in hand with early or later drop-out. The analysis of additional cognitive, motivational, and external variables gives clues for instructional interventions. However, a lot of limitations, challenges and open questions remain. First, the number of participants in our study is far too low to yield valid and reliable results, which could be transferred to other cohorts of students. Only a validation of our prediction model with the next student cohort in the course can reveal the quality of our results. Additionally, the instructional interventions are only derived from data analysis. These aspects require further studies. Finally, yet importantly, the prediction model acts on the assumption that each student works on her own while solving the exercise. If students work together, as it is quite common in computer science and could be hardly prevented in online exercises, there might be effects on the programming behavior as well as for the decision for drop-out or persistence. In further studies, the possibility of cooperative programming has to be reflected in the design of the study.

Note 1. The CSharp is hosted at: https://github.com/liaScript/Csharpcourse

References


### Appendix 1. Evaluated variables: log file

<table>
<thead>
<tr>
<th>Context Variable</th>
<th>Meaning</th>
<th>Data type</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Programming style</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>action_count</td>
<td>Mean number of code lines changed per commit</td>
<td>Integer</td>
</tr>
<tr>
<td>Tn_mean_dLOCC</td>
<td>Mean number of code lines changed per commit</td>
<td>Integer</td>
</tr>
<tr>
<td>Tn_max_dLOC</td>
<td>Max number of code lines changed per commit</td>
<td>Integer</td>
</tr>
<tr>
<td>Tn_mean_dTime</td>
<td>Mean duration between two commits (saturation at 10 min)</td>
<td>Float</td>
</tr>
<tr>
<td>Tn_var_dTime</td>
<td>Variance of the duration between two commits (min²)</td>
<td>Float</td>
</tr>
<tr>
<td><strong>Errors and debugging</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>success_count</td>
<td>Count of successful commits</td>
<td>Integer</td>
</tr>
<tr>
<td>success_ratio</td>
<td>Ratio of successful commits compared to action_counts</td>
<td>Float</td>
</tr>
<tr>
<td>error_count, fatal_count, warning_count</td>
<td>Counts of specific error classes</td>
<td>Integer</td>
</tr>
<tr>
<td>error_ratio, fatal_ratio, warning_ratio</td>
<td>Ratio of the specific error class compared to action_counts</td>
<td>Float</td>
</tr>
<tr>
<td>X_count</td>
<td>Count of error type X (125 error types according to Compiler outputs)</td>
<td>Float</td>
</tr>
<tr>
<td>X_ratio</td>
<td>Ratio of error type X related to error count</td>
<td>Float</td>
</tr>
<tr>
<td>S_count</td>
<td>Error streak count</td>
<td>Integer</td>
</tr>
<tr>
<td>S_ratio</td>
<td>Ratio of streaks compared to action_counts</td>
<td>Float</td>
</tr>
<tr>
<td>S_trans_max</td>
<td>Max length of a streak</td>
<td>Integer</td>
</tr>
<tr>
<td>S_duration_max</td>
<td>Max duration to cope with a streak (saturation at 10 min)</td>
<td>Float</td>
</tr>
<tr>
<td>S_trans_mean</td>
<td>Mean length of a streak</td>
<td>Float</td>
</tr>
<tr>
<td><strong>External variable</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tn_max_delay</td>
<td>Maximum delay between publication of a task and first commit (days)</td>
<td>Float</td>
</tr>
<tr>
<td><strong>Results</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Passed</td>
<td>Whole course successfully finished</td>
<td>Boolean</td>
</tr>
<tr>
<td>Tn_finished</td>
<td>Individual task n finalized</td>
<td>Boolean</td>
</tr>
</tbody>
</table>

*Table AI.* Identified variables from log for user behavior
Appendix 2. Questionnaire

<table>
<thead>
<tr>
<th>Context</th>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>User: demographic data</td>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Year of birth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Study course</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Term</td>
<td></td>
</tr>
<tr>
<td>User: self-assessment</td>
<td>Time spent</td>
<td>Time on-campus in lecture</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Time for self-study</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Time for work (student job)</td>
</tr>
<tr>
<td></td>
<td>Programming experiences</td>
<td>For how many years have you been programming?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>How experienced are you with the following languages?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Assembler, C, C++, Java, Python</td>
</tr>
<tr>
<td></td>
<td>Perceived programming skills</td>
<td>Self-estimation: on a scale from 1 to 5, how do you estimate your</td>
</tr>
<tr>
<td></td>
<td></td>
<td>programming skills?</td>
</tr>
<tr>
<td></td>
<td>Self-efficacy</td>
<td>I am confident about my programming knowledge</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I am confident that I can apply the programming skills in solving problems</td>
</tr>
<tr>
<td></td>
<td></td>
<td>In general, I could complete any desired programming task also if […]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[…] there was no one around to tell me what to do</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[…] I had never solved a task like it before</td>
</tr>
</tbody>
</table>

| Table AII. Variables from questionnaire | Prior knowledge questionnaire | 22 multiple choice questions for prior knowledge classification on the topic of the lecture and two questions with code snippets and free text answers |