

# Small data as a conversation starter for learning analytics

## Exam results dashboard for first-year students in higher education

Tom Broos

*Department of Computer Science, Katholieke Universiteit Leuven, Leuven, Belgium and*

*Department of Tutorial services, Faculty of Engineering Science, Katholieke Universiteit Leuven, Leuven, Belgium*

Katrien Verbert

*Department of Computer Science, Faculty of Engineering Science, Katholieke Universiteit Leuven, Leuven, Belgium*

Greet Langie

*Faculty of Engineering Technology, Katholieke Universiteit Leuven, Leuven, Belgium*

Carolien Van Soom

*Faculty of Science, Katholieke Universiteit Leuven, Leuven, Belgium, and*

Tinne De Laet

*Faculty of Engineering Science, Katholieke Universiteit Leuven, Leuven, Belgium*

### Abstract

**Purpose** – The purpose of this paper is to draw attention to the potential of “small data” to complement research in learning analytics (LA) and to share some of the insights learned from this approach.

**Design/methodology/approach** – This study demonstrates an approach inspired by design science research, making a dashboard available to  $n = 1,905$  students in 11 study programs (used by  $n = 887$ ) to learn how it is being used and to gather student feedback.

**Findings** – Students react positively to the LA dashboard, but usage and feedback differ depending on study success.

**Research limitations/implications** – More research is needed to explore the expectations of a high-performing student with regards to LA dashboards.

**Originality/value** – This publication demonstrates how a small data approach to LA contributes to building a better understanding.

**Keywords** Higher education, Case study, First-year students, Learning analytics, Dashboard, Small data

**Paper type** Research paper

### Introduction

Learning analytics (LA) is a growing domain of research that is on the one hand enthusiastically exploring the potential of big data, but on the other hand provoking more and more critical considerations, for instance about ethics and privacy. Referring to maturity models in

© Tom Broos, Katrien Verbert, Greet Langie, Carolien Van Soom and Tinne De Laet. Published in the *Journal of Research in Innovative Teaching & Learning*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

This research is co-funded by the Erasmus + program of the European Union (STELA Project, ref. 562167-EPP-1-2015-1-BE-EPPKA3-PI-FORWARD).



---

information systems, we argue that LA research involving big data can be complemented by a more modest approach based on “small data” that is readily available within institutions. This approach makes it possible to learn about expectations regarding, and challenges for LA within the short term. We demonstrate this approach by presenting an evaluation study with  $n = 1,905$  first-year students, offering them a dashboard for their exam results.

### *A plea for small data and modest LA.*

LA is defined as “the measurement, collection, analysis and reporting of data about learners and their context, for purposes of understanding and optimizing learning and the environments in which it occurs.” (Siemens and Long, 2011). In contrast to the related concept of academic analytics, its primary subject is the learning process, while the latter has a broader scope. In this perspective, learners and faculty are the primary beneficiaries of LA, while administrators, governments, and authorities are exemplary for the target audience of academic analytics (Siemens and Long, 2011). Some insist on a further distinction, equating academic analytics to a concept parallel to business analytics, supporting operational, and financial (business) decision making, but in the specific context of higher education institutions (Van Barneveld *et al.*, 2012).

In turn, the term analytics became a nominator for techniques applied when analyzing big data (Gandomi and Haider, 2015). Big data is different from “small” data due to its volume (terabytes, petabytes), its velocity (real-time or near real-time), and variety (including both structured and unstructured data) (Laney, 2001). It is exhaustive and aims to capture entire populations ( $n = \text{all}$ ) (Kitchin, 2013), which further elucidates the distinction between big data analytics and traditional statistical analysis that regularly involves dealing with samples limited in size.

To some authors, big data analytics preludes “the end of theory,” making not just *ex ante* hypothesis unnecessary – if there is a pattern in the data, the algorithms will find it anyway – but even the need to theorize *ex post* about underlying causation or explanations (Prensky, 2009): “Forget taxonomy, ontology, and psychology. Who knows why people do what they do? The point is they do it, and we can track and measure it with unprecedented fidelity. With enough data, the numbers speak for themselves” (Anderson, 2008). It may come to no surprise that many scientists feel uncomfortable about this notion (Kitchin, 2013; Lazer *et al.*, 2014) and that it may contribute to a distrust of LA from educational scientists. In his book “The Black Box Society,” Pasquale (2015) portrays a dystopian view on the growing use of big data to drive all kinds of decisions. He sketches algorithms that are unfair or unaccountable, proprietary, complex, and self-fulfilling.

Apart from the analytics, the collection of massive amounts of personal data is subject of debate in itself. One concern is, of course, the level of privacy. Because of the fine-grained character of big data, several examples of re-identification attacks demonstrated the difficulty of giving any guarantees about anonymization (Boyd and Crawford, 2012; Jensen, 2013). Ever increasing computational power may make data collection that is respecting today’s privacy requirements entirely retraceable to the personal level within a few years. Attempts to make the de-identification process more robust may diminish the utility of the data (Daries *et al.*, 2014). Again, it may not come unexpected that many, not in the least data subjects themselves, such as teachers, students, and their parents, are reluctant to the use of big data in education. In 2014, the inBloom project fell victim to “exaggerated fears and a misunderstanding about the technology” (C.K.N., 2014). The project was conceived by American educators to answer a shared problem of managing student data in a secure, standardized, and efficient way, but its failure to convince concerned parents and privacy advocates resulted in a shutdown.

Within the European legal framework, the growing scope of data collection in education should be subjected to the proportionality principle (Berendt *et al.*, 2017): the data should: serve a legitimate aim within the broader view of the role of education in society; the use of

data should be suitable to achieve the aim, as demonstrated by rigorously constructed scientific evidence; the data should be necessary, in the sense that similar results cannot be obtained with less; its collection should be reasonable, meaning that the objectives are balanced against the loss of privacy.

Privacy became an increasingly discussed topic in the domain of LA (e.g. Gašević *et al.*, 2016; Hoel *et al.*, 2017; Prinsloo and Slade, 2017), not only as an obstruction but as a constraint that needs to be considered early in the design. At the same time, efforts are being made to close the gap between “traditional” research in education and the use of big data (e.g. Gibson and Ifenthaler, 2017).

In their book “Competing on Analytics,” Davenport and Harris (2007) describe an analytics maturity model with five stages ranging from the “analytically impaired” organization that collects data but is generally incapable of deriving actionable insights from it, to “analytical competitors” that are entirely built around analytics. Maturity assessment models are common practice in management science and information systems. They provide a way to determine and describe the current state of organization within a specific scope and allow for goal-setting and gap analysis. Maturity can be evaluated from different perspectives (Mettler, 2011): process/structure, object/technology, and people/culture. Maturity or sophistication models describe a growth path, hence skipping maturity stages “seldom proves to be wise. There is valuable experience to be gained at each stage [...]” (Wixom and Watson, 2012). Siemens *et al.* (2013) proposed an “LA Sophistication Model” that starts out from simple awareness and gradually shifts into the direction of transformation at the sector level.

This paper aims to complement the big data approach to LA, a promising but also challenging endeavor, with a more “modest” implementation in practice. A less esoteric approach facilitates an immediate involvement of students, practitioners, managers, and policymakers. Each of these stakeholders can offer useful input about their expectations, experiences, and suggestions for improvement. Apart from providing valuable input for further research, this approach contributes to a gradual introduction of LA, managing change in line with the principles addressed by maturity models. This is not to say that researchers should abandon big data in LA altogether, to the contrary. Big data has a long road ahead toward general acceptance in education. It will take time before questions about proof of value are sufficiently answered to justify the collection of vast amounts of personal data for LA. Meanwhile, there is a lot to be gained from bringing more conservative applications of LA into practice.

#### *Involving students and study counselors*

Designing LA interventions is not limited to technological aspects, but includes the entire educational context and the processes surrounding it (Wise, 2014; Gašević *et al.*, 2015). An increasing number of authors are taking a critical position toward LA (e.g. Roberts *et al.*, 2017), some addressing the lack of knowledge about the topic among students and the need to engage students in the decision making (Roberts *et al.*, 2017). And indeed, a lot can be learned by engaging students and practitioners in focus groups to find out about their concerns and suggestions. In this paper, we discuss an alternative route inspired by design science research (DSR) in information systems – “learning through the act of building” (Kuechler and Vaishnavi, 2008). Many things can be learned by exposing a sizable number of students to iterative implementations of learning dashboards early in the process. This ensures that the basic concepts and subsequent development of LA within the institution become conceivable and provides handlebars for further discussion.

Several types of data about students are already being collected by higher education institutions. The data are being used to support research, management, and policy, but in several (most) cases not being fed back to students in an insightful way, if at all. Doing so

---

now in the context of LA provides an opportunity to start the conversation about LA in a non-threatening manner and a gentle way to guide the organization into the direction of an ongoing sophistication process.

### *Background*

We have put the concept of starting with small data interventions into practice. In the next session, we present the MyScores dashboard that was offered to 1905 first-year students in 11 science and engineering (STEM) oriented study programs within the University of Leuven (KU Leuven) following shortly after the publication of their exam scores.

Within the different faculties at KU Leuven, study counselors are available to support first-year students in their transition from secondary to higher education. Much experience is present in these study counseling services. Study counselors are bridge persons; their role involves the reconciliation of interests from faculty and students and we consider them well-placed advisors when it comes to acceptance of LA interventions. By involving them actively in a co-creation process – for example, by asking them to contribute the accompanying texts and to evaluate the visualizations in the dashboard discussed below – some of the experience can be captured and a friendly communication line is opened to inform further research on LA, including studies that do have big data in scope.

### *MyScores dashboard*

In this section and the next, we use the categories of Bodily and Verbert (2017) to introduce the MyGrades student dashboard (Figure 1) in a structured and comparable manner (see Table I for an overview).

*Intended goal.* The goal of the MyScores dashboard was to support students in the first year of higher education by providing them with the feedback on exam results, encouraging reflection and adding actionable recommendations for improvement. The dashboard was to demonstrate how simple, “small data” that is readily available within institutions can be used as a starting point to fuel the organizational learning process around LA and to collect the valuable feedback about the intervention.

*Information selection.* The MyScores dashboard visualizes exam results on the level of individual courses and the study program. These have been stored in computer systems for decades and results are being communicated to students in an electronic format for many years at KU Leuven, but the traditional “study progress file” only offers a limited and unprepossessing presentation.

*Needs assessment.* The first examination period in Flanders takes place in the middle of the academic year. On the course level, several reasons may lead to a certain degree of ambiguity when newcomers in higher education receive their first exam report. Students are not “graded on a curve,” meaning that scores are not relative to those of other students. Belgian universities use a common rating scale from 0 to 20, with 10 being the passing grade. In many study programs, it is common for distributions of scores to differ significantly from one course to another. Most secondary schools use different grading scales and even in relative terms, the comparability of results in secondary and higher education is limited. On the program level, students need to get acquainted with a crucial key performance indicator: the cumulative study efficiency (CSE). This is the ratio of credits a student passed throughout the years to the total number of credits taken within the study program. Most programs have a semester total of 30 credits and typical one-semester courses correspond to three to six credits. The CSE is an important indicator of progress and connected to institutional restrictions. Students who do not attain a CSE of at least 30 percent at the end of the first year are not allowed to continue the program; students with

The screenshot shows a dashboard with five tabs: 'Inleiding', 'Scores per vak' (selected), 'Globaal', 'Tips', and 'Regelgeving'. The main content area is divided into sections for different courses, each with a score and a feedback prompt.

**Inleiding** | **Scores per vak** | **Globaal** | **Tips** | **Regelgeving**

Hier vind je een zelfreflectieinstrument waarbij je voor de eerstejaarsvakken van ingenieurswetenschappen aangeeft hoe je inschatting was na het examen en hoe tevreden je bent met je resultaat. Zodra je deze reflectie hebt gemaakt kan je er voor kiezen om je resultaat te positioneren ten opzichte van andere studenten die dit vak binnen een diplomacontract geboekt hebben in hun ISP. In de visualisatie stelt elk (te)groen, groen, oranje of rood bolletje één student voor.

Let op: dit instrument is gemaakt om zelfreflectie te initiëren. Nadenken over je resultaten, is louter de start van een zelf-reflectieproces. Je resultaten tonen uiteraard maar een fractie van wie jij bent als student. Dit platform kan onmogelijk rekening houden met persoonlijke factoren en omstandigheden die mogelijk je studie beïnvloeden. Jij bent zelf het beste geplaatst om de andere factoren die belangrijk zijn in rekening te brengen.

**ANALYSE, DEEL 1**  
Jouw score: 12/20

- Deze score is (kies een optie) dan wat ik zelf na het examen verwachtte.
- Ik ben (kies een optie) over dit resultaat.

Toon de scores van jaargenoten

---

**TOEGEPASTE ALGEBRA**  
Jouw score: 15/20

- Deze score is (kies een optie) dan wat ik zelf na het examen verwachtte.
- Ik ben (kies een optie) over dit resultaat.

Toon de scores van jaargenoten

---

**ALGEMENE EN TECHNISCHE SCHEIKUNDE**  
Jouw score: 9/20

- Deze score is (kies een optie) dan wat ik zelf na het examen verwachtte.
- Ik ben (kies een optie) over dit resultaat.

Toon de scores van jaargenoten

---

**TOEGEPASTE MECHANICA, DEEL 1**  
Jouw score: 10/20

- Deze score is (kies een optie) dan wat ik zelf na het examen verwachtte.
- Ik ben (kies een optie) over dit resultaat.

Toon de scores van jaargenoten

---

**PROBLEEMOPLOSSEN EN ONTWERPEN, DEEL 1**  
Jouw score: 15/20

**Jouw mening over deze informatie zou ons erg helpen!**

(-) 1 2 3 4 5 (++)  
Ik vind deze informatie nuttig.

(-) 1 2 3 4 5 (++)  
Ik vind deze informatie duidelijk.

(-) 1 2 3 4 5 (++)  
Deze info heeft invloed op hoe tevreden ik ben over mijn resultaten.

**Figure 1.** Initial view of the “Scores per course” tab, before answering questions about expectations and satisfaction

**Notes:** The text, in Dutch, at the top explains the goal of the current tab (self-reflection), its contents and how to read the visualizations. The second paragraph reminds the student to the fact that the dashboard is only a partial representation, not accounting for individual circumstances. At the bottom of the dashboard, students are being invited to provide feedback on the dashboard

a CSE below 50 percent receive binding study advice and may eventually be refused enrollment to all programs if they fail to catch up.

First-year students may find it difficult to properly assess their situation. In the presence of such uncertainty, social-comparison theory (Festinger, 1954) suggests that

individuals will try to compare themselves to peers. The MyScores dashboard responds to this strategy, by offering students the option to compare their exam scores to those of other participants and to position their CSE in relation to other students in the same study program.

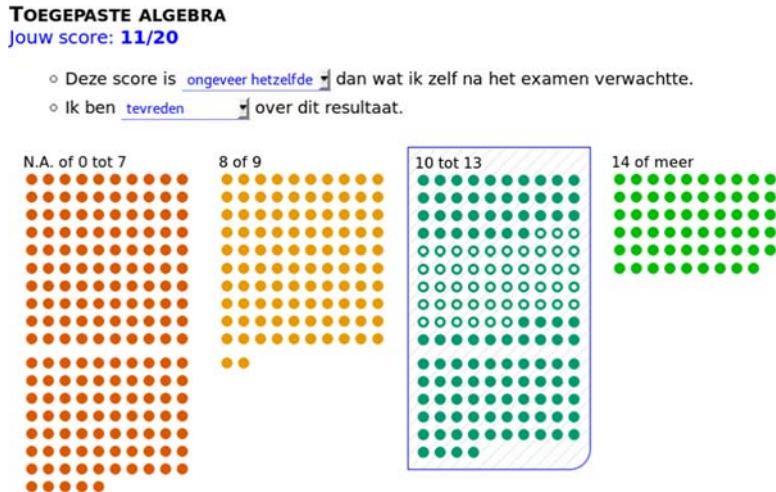
Additionally, the dashboard includes historic data that connects the first-year, first-semester CSE students from previous years to their success rate and time needed to finish the program. The objective is to help students to become aware of the meaning and impact of the progress indicator, beyond the simple number. Previously, only one of the participating faculties offered such guidance in a systematic manner, but without further personalization of the message.

*Visual design.* Most student dashboards represent information in a visual way, but it has been argued that including elaborate textual information is useful to provide additional guidance (Ramos-Soto *et al.*, 2015). Previous experience within the institution (Broos *et al.*, 2017) reinforced the idea that the provision of a flexible text parameterization system for study counselors is a good way to make use of their knowledge and experience and to foster their involvement in and acceptance of the intervention. Hence, in addition to data visualizations discussed below, the dashboard contains a considerable number of accompanying notes that are a composition of common parts, parts adapted to the study program and parts that are individualized based on the actual situation and grades of the student. To offer a better overview, the dashboard is composed out of five tabs:

- (1) Introduction: the first tab contains an introductory text, explaining the objectives, and contents of the other four tabs. It includes a message explaining that scores are only a partial view of who the students are and provide a direct link to the study counselor and/or support service assigned to the study program.
- (2) Scores per course: the second tab contains the exam results for each course the student is enrolled in. The initial view shows a numeric score on a 0-20 scale and prompts the student to select the applicable option for two statements: “This score is [...] than what I expected after the exam” and “I feel [...] about this result.” When both blanks are filled, a button to show the scores of peer students is unlocked. On pressing this button, a visualization containing the scores of all exam participants appears. Within this visualization, the student is positioned in one of four groups. A second button offers the option to position the individual score more precisely within the group, leaving it up to the students to choose whether they would like to see this information or not (see Figure 2). This sequence is repeated for each of the enlisted courses. The tab concludes with an invitation to reflect on what is shown above, some additional remarks to frame what is shown above and an open invitation to get in touch with a student counselor.

Category	Question
Intended goal	What is the intended goal of the system?
Information selection	What types of data support your goal?
Needs assessment	What do students need? Does this need align with your goal?
Visual design	Why are you using the visual techniques or recommendations you have chosen?
Visualizations	What visual techniques will best represent your data?
Student use	How are students using the system? How often? Why?
Actual effects	What is the effect on student behavior/achievement?
Student perceptions	How do students perceive the reporting system?
Usability test	Is the system easy and intuitive to use? Usability

**Table I.**  
Categories used to present the MyScores dashboard, adapted from “Questions to guide in implementing reporting systems” (Bodily and Verbert, 2017)



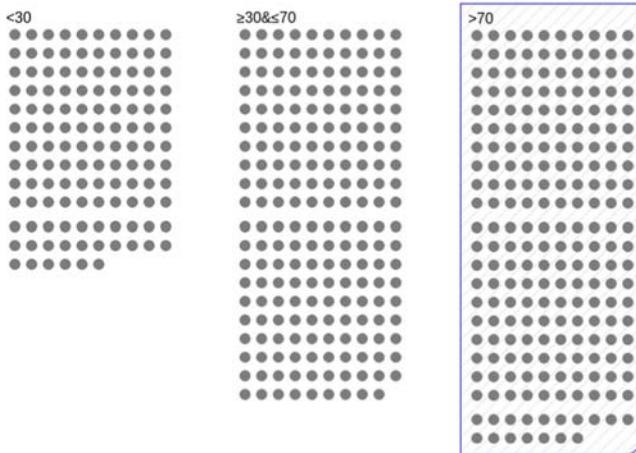
**Notes:** The group that applies to the student is marked with a blue border and hatched background pattern. After pressing the button to receive a more precise positioning, dots of students with the same score are marked differently

**Figure 2.**  
Expanded  
visualization  
for one course

- (3) Global progress: the middle tab informs about the CSE. The upper part (see Figure 3) summarizes the progress of the student on the program level. It contains a visualization to put the individual result in perspective, comparing it to other students within the same study program. As with the course level scores, the initial view is intentionally coarse. When preferred, a button allows the student to activate a more fine-grained positioning. The lower part (see Figure 4) of the “global” tab explains and visualizes how the CSE of students within the same study program in previous years relates the number of years they spent to obtain the bachelor diploma: three years (default), four years or five or more. Drop-out students are indicated as well. For the category that corresponds best to the student, this information is repeated in an accompanying text.
- (4) Tips: the fourth tab’s content is entirely textual and is divided into two parts. The first part contains four sections on how to process the information contained by the dashboard: “talk about it,” “realistic expectations,” “learning skills” – including a link to our previous dashboard on learning skills (Broos *et al.*, 2017) – and other factors (e.g. personal situation). The second part contains actionable advice on how to set goals and tips on how to achieve them.
- (5) Regulations: the last tab provides a summary of the progress monitoring regulations with respect to the CSE (see above). As with most of the textual information in the dashboard, this information is also available online, but it is fragmented across several web pages following an organization oriented structure. The MyScores dashboard collects relevant information from different sources and offers an integrated and personalized view, adapted to the program and situation of the student, and focused on the specific context of the first exam results.

*Visualization.* The dashboard uses a consistent visualization method across different measures. Unit charts are used to assign student scores to groups. Every student is represented by one dot and attributed to one of the groups. In case of the course scores, the dots are color-coded according to their impact on the student: red for failed and

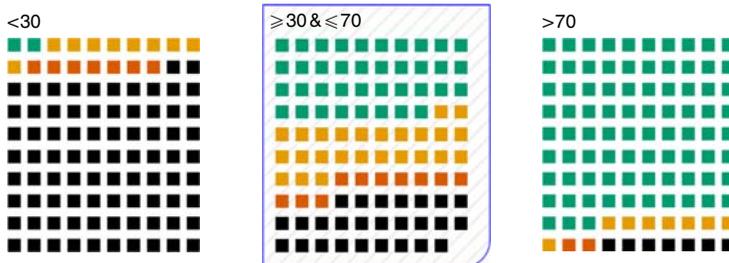
De visualisatie hieronder toont je positionering op basis van je behaalde studie-efficiëntie (CSE) van 73% ten opzichte van de andere eerstejaarsstudenten ingenieurswetenschappen van dit academiejaar. Elk grijs bolletje stelt één student voor. De categorie waarin jij je op basis van je studie-efficiëntie bevindt, **jouw groep**, is nog eens expliciet aangeduid.



Positioneer mijn score op deze grafiek.

**Notes:** Top: example of text that accompanies the visualization. Bottom: button to activate a more precise positioning if desired by the student

**Figure 3.** Visualization of the CSE for current year's students, divided into three groups



**Note:** The squares do not represent individual students, but percentage units

**Figure 4.** CSE visualization for previous year's students

non-tolerable (0 to 7/20) or not completed, orange for failed but with a potentially tolerable score (8 or 9/20), green for passed with moderate score (10 to 13/20), and bright green for passed with high score (14 to 20/20). For the global process tab, the first visualization (see Figure 3) uses gray dots for current year's students, combined into three groups (depending on the study program, e.g. 0-30, 30-70, and 70-100 percent for engineering science). The second visualization (see Figure 4) uses colored squares to represent percentage units of previous year's students (rather than dots for single students). Here, colors correspond to the time these students required to obtain the bachelor degree: black for students that dropped out, red for those who needed five years or more, orange for students with one year of delay (four years), and green for students that obtained the degree within the default three years. For each of the visualizations, the interpretation of colors and markers (dot or square) is explained using inline elements within the text.

If students activate a more precise positioning of their own result (by clicking the corresponding button) the dots of all students with the same score become emphasized (unfilled and blinking a few times to catch attention).

**Results and discussion**

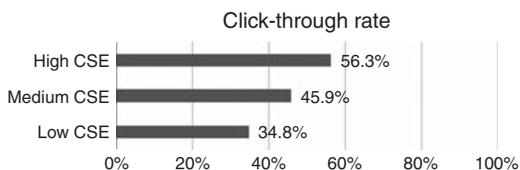
*Student use*

Out of 1,905 unique students that received an invitation by e-mail, 887 (47 percent) used the dashboard. There is a remarkable difference in click-through rate between study programs. Students in programs bioengineering, engineering science, and engineering science-architecture received the invitation one day after online publication of their results in the university’s portal. In total, 55, 68, and 60 percent of students in these programs clicked through to the dashboard, respectively. For students of the science faculty (chemistry, biology, biochemistry-biotechnology, geography, geology, mathematics, informatics and physics) a different procedure applies. Students need to collect results in person allowing study counselors to provide some extra information on the spot. These students received the dashboard invitation only after this procedure was completed, which may partly explain why the click-through rates were only between 22 percent and 40 percent for these programs.

There is a noteworthy difference in the click-through rate between CSE groups (see Figure 5). More than half (56.3 percent) of the students with a relatively good study progress – the high CSE group[1] – do click through from invitation to dashboard. Among students with an alarming low study progress – the low CSE group – only 34.8 percent are visiting the dashboard. The medium CSE group has a click-through rate of 45.9 percent. Possibly, this difference may be explained by students who have already given up and may have lost interest in further information. Additionally, it is possible that student characteristics that result in a lower interest in the dashboard also result in lower exam results. Although more research is required to explain the exact cause of the difference, the suggestion that student dashboards are more likely to appeal to “successful” students is already informative, especially when considering that many efforts in LA are focused on at-risk students and drop-out prevention. On the one hand, the suggestion is that additional effort is required to reach “weaker” students. On the other hand, dashboards should be adapted to provide appropriate feedback that is valuable to “successful” students, for example by informing them about honor programs or outlining the prospect of advanced academic training.

*Actual effects*

Verbert *et al.* (2013) presented a conceptual framework for evaluating LA applications that distinguishes four stages: awareness, reflection, sense-making, and impact. The awareness stage is limited to simply displaying data. Almost by definition, any dashboard would



**Notes:** The width of the rectangles is an indication of the CSE distribution for all invited students. The relative heights correspond to the ratio of students within the CSE group that clicked through to the dashboard

**Figure 5.** Click-through rate in relation to the study progress of the student

achieve this phase. But the presentation of data is not the end goal. The reflection stage relates to how the data are used to provoke questions in the mind of the user. In the MyScores dashboard, we opted to steer students in this direction by requiring them to answer questions related to their self-assessment and satisfaction to unlock more detailed information. By placing individual results within context, the MyScores dashboard aims to provoke additional questions, which in case of students' endeavor to answer them, should lead to new insights in the sense-making phase.

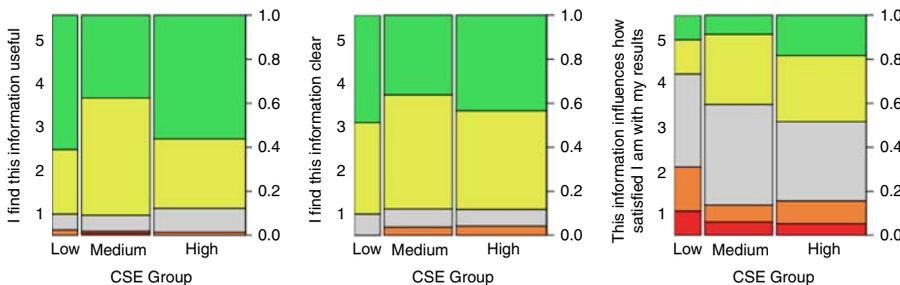
The fourth phase of the framework, impact, requires the assessment of behavioral change. While the MyScores dashboard provides several starting points to guide the students in the direction, the assessment of impact would require a longitudinal study and was left out of scope here. To assess awareness, reflection, and sense-making, three perspectives were used: do students use the footholds offered by the dashboard, how do students perceive the dashboard, and does the evaluation of the dashboard depend on the situation (grades, progress) of the student?

*Student perceptions and usability*

Students were asked to respond to three statements on a 1 (–) to 5 (+) scale:

- (1) I find this information useful (usefulness);
- (2) I find this information clear (clearness); and
- (3) this information influences how satisfied I am with my results (reflection).

These statements were presented to them in a noticeable yellow area at the bottom of the dashboard (see bottom Figure 1). Results are summarized in the Marimekko charts, as shown in Figure 6. In total, 289 out of 887 (32.6 percent) dashboard users provided full feedback on all three statements. Usefulness was rated positively (response 4-5) by 87.7 percent of feedback-providing students within the high CSE group for their study program, 90.8 percent by the medium CSE group, and 90.3 percent by the low CSE group. Clearness was rated similarly: 88.2, 88.0, and 90.3 percent in the high, medium, and low CSE groups, respectively. Students within the medium CSE group seem to respond with a slightly lower score on average for both statements. When asked if the information provided by the dashboard has influence on how satisfied students are with their results, based on the idea that providing perspective may alter how students think and feel about their own results, the response is more moderate. In total, 48.4 percent of students in the high CSE group, 40.5 percent of students in the medium CSE group, and 26.7 percent of students in the low CSE group provide a positive response. A large share of students provides an answer somewhere in the middle (35.9, 45.7 and 42.2 percent for the three groups, respectively). Up to 31.1 percent of students in the low CSE group seem to indicate that the dashboard did not alter satisfaction with their results. In retrospect, the third



**Note:** From left to right: usefulness, clearness and reflection

**Figure 6.** Marimekko charts encoding the responses to the three statements

statement may have given rise to some misinterpretation: some students may have intended to indicate that the information presented influenced their satisfaction negatively, rather than not (see Figure 1, lower right).

### Evaluation and conclusion

This paper suggests complementing big data for LA within higher education institutions by a more modest, practice-oriented approach leveraging existing small data and to start gathering additional insights by deploying dashboards to students. Subsequently, it presents an implementation of this approach by offering the MyScores dashboard to 1905 first-year students in 11 science and engineering-oriented study programs, reaching 887 active users, from which 289 provided feedback on three statements about the dashboard. Generally, students rate the usefulness and clearness of the MyScores dashboard positively, but the appreciation of how the dashboard changes their satisfaction about their results is less consistent and differs depending on the study success. Moreover, “weaker” students seem to access the dashboard less, which raises questions about how to reach these students, but also underline the potential of targeting “stronger” students with targeted information. The difference in dashboard access between “weaker” and “strong” students is consistent with earlier finding within a largely overlapping population of STEM students (Broos *et al.*, 2017). Further research is required to explore if “weaker” students are less likely to use LA tools overall, or if the finding is specific for the type of dashboards or the process used to inform students about their availability. For a follow-up study, we would choose a more structured approach, based on DSR, using dashboard artifacts to further probe into how students accept and use them, complementing qualitative usage information and questionnaire feedback by in-depth qualitative research. Focus group discussions may help to build the understanding of the difference in usage and response between students with low, medium and high study efficiency (CSE).

The practical implementation of a dashboard targeting a sizable number of first-year students involved many contributors, including students, study counselors, university and faculty management, IT and legal services, etc. This allowed us to see LA at the organizational level in its full complexity and from many angles, some of them previously overlooked. Many of the insights gained throughout the process were not discussed in this paper.

Our study remained on the surface with respect to the lessons learned from the MyScores intervention and would ideally be complemented by a more longitudinal approach to try to capture is the dashboard leads to behavioral change (impact). For each of the participating STEM study programs we received confirmation for continued participation. Several additional study programs confirmed participation in an upcoming iteration, including programs in humanities, social and biomedical sciences, expanding the population both in size and diversity.

### Note

1. The allocation of students to a high, medium or low CSE group is based on program-specific CSE cutoff scores. The analysis preserves the allocation as used on the dashboard.

### References

- Anderson, C. (2008), “The end of theory: the data deluge makes the scientific method obsolete”, *Wired Magazine*, Vol. 16 No. 7, available at: [www.wired.com/2008/06/pb-theory](http://www.wired.com/2008/06/pb-theory) (accessed July 28, 2017).
- Berendt, B., Littlejohn, A., Kern, P., Mitros, P., Shacklock, X. and Blakemore, M. (2017), “Big Data for monitoring educational systems”, Publications Office of the European Union.
- Bodily, R. and Verbert, K. (2017), “Trends and issues in student-facing learning analytics reporting systems research”, *Proceedings of the 7th International Learning Analytics & Knowledge Conference, March*, pp. 309-318.

- Boyd, D. and Crawford, K. (2012), "Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon", *Information, Communication & Society*, Vol. 15 No. 5, pp. 662-679.
- Broos, T., Peeters, L., Verbert, K., Van Soom, C., Langie, G. and De Laet, T. (2017), "Dashboard for actionable feedback on learning skills: scalability and usefulness", in Zaphiris, P. and Ioannou, A. (Eds), *Proceedings of the 4th International Conference on Learning and Collaboration Technologies*, Springer, Cham, pp. 229-241.
- C.K.N. (2014), "Big data and education: withered in bloom", *The Economist*, April, available at: [www.economist.com/blogs/schumpeter/2014/04/big-data-and-education/](http://www.economist.com/blogs/schumpeter/2014/04/big-data-and-education/) (accessed July 28, 2017).
- Daries, J.P., Reich, J., Waldo, J., Young, E.M., Whittinghill, J., Ho, A.D., Seaton, D.T. and Chuang, I. (2014), "Privacy, anonymity, and big data in the social sciences", *Communications of the ACM*, Vol. 57 No. 9, pp. 56-63.
- Davenport, T.H. and Harris, J.G. (2007), *Competing on Analytics: The New Science of Winning*, Harvard Business Press, Boston, MA.
- Festinger, L. (1954), "A theory of social comparison processes", *Human Relations*, Vol. 7 No. 2, pp. 117-140.
- Gandomi, A. and Haider, M. (2015), "Beyond the hype: big data concepts, methods, and analytics", *International Journal of Information Management*, Vol. 35 No. 2, pp. 137-144.
- Gašević, D., Dawson, S. and Jovanovic, J. (2016), "Ethics and privacy as enablers of learning analytics", *Journal of Learning Analytics*, Vol. 3 No. 1, pp. 1-4.
- Gašević, D., Dawson, S. and Siemens, G. (2015), "Let's not forget: learning analytics are about learning", *TechTrends*, Vol. 59 No. 1, pp. 64-71.
- Gibson, D.C. and Ifenthaler, D. (2017), "Preparing the next generation of education researchers for big data in higher education", in Kei Daniel, B. (Ed.), *Big Data and Learning Analytics in Higher Education*, Springer, Cham, pp. 29-42.
- Hoel, T., Griffiths, D. and Chen, W. (2017), "The influence of data protection and privacy frameworks on the design of learning analytics systems", *LAK, March*, pp. 243-252.
- Jensen, M. (2013), "Challenges of privacy protection in big data analytics", *IEEE International Congress on Big Data, June*, pp. 235-238.
- Kitchin, R. (2013), "Big data and human geography: opportunities, challenges and risks", *Dialogues in Human Geography*, Vol. 3 No. 3, pp. 262-267.
- Kuechler, B. and Vaishnavi, V. (2008), "On theory development in design science research: anatomy of a research project", *European Journal of Information Systems*, Vol. 17 No. 5, pp. 489-504.
- Laney, D. (2001), "3D data management: controlling data volume, velocity and variety", *META Group Research Note*, Vol. 6, p. 70.
- Lazer, D., Kennedy, R., King, G. and Vespignani, A. (2014), "The parable of Google Flu: traps in big data analysis", *Science*, Vol. 343 No. 6176, pp. 1203-1205.
- Mettler, T. (2011), "Maturity assessment models: a design science research approach", *International Journal of Society Systems Science*, Vol. 3 Nos 1-2, pp. 81-98.
- Pasquale, F. (2015), "The black box society: the secret algorithms that control money and information", Harvard University Press.
- Prensky, M. (2009), "H. Sapiens digital: from digital immigrants and digital natives to digital wisdom", *Innovate: Journal of Online Education*, Vol. 5 No. 3, Article No. 1.
- Prinsloo, P. and Slade, S. (2017), "Big data, higher education and learning analytics: beyond justice, towards an ethics of care", in Kei Daniel, B. (Ed.), *Big Data and Learning Analytics in Higher Education*, Springer, Cham, pp. 109-124.
- Ramos-Soto, A., Lama, M., Vázquez-Barreiros, B., Bugarín, A., Mucientes, M. and Barro, S. (2015), "Towards textual reporting in learning analytics dashboards", *2015 IEEE 15th International Conference on Advanced Learning Technologies, July*, pp. 260-264.

- Roberts, L.D., Chang, V. and Gibson, D. (2017), "Ethical considerations in adopting a university-and system-wide approach to data and learning analytics", in Kei Daniel, B. (Ed.), *Big Data and Learning Analytics in Higher Education*, Springer, Cham, pp. 89-108.
- Siemens, G. and Long, P. (2011), "Penetrating the fog: analytics in learning and education", *EDUCAUSE Review*, Vol. 46 No. 5, pp. 30-40.
- Siemens, G., Dawson, S. and Lynch, G. (2013), *Improving the Quality and Productivity of the Higher Education Sector. Policy and Strategy for Systems-Level Deployment of Learning Analytics*, ACT: Society for Learning Analytics Research for the Australian Office for Learning and Teaching, Canberra.
- Van Barneveld, A., Arnold, K.E. and Campbell, J.P. (2012), "Analytics in higher education: establishing a common language", *EDUCAUSE Learning Initiative*, Vol. 1 No. 1, pp. 1-11.
- Verbert, K., Duval, E., Klerkx, J., Govaerts, S. and Santos, J.L. (2013), "Learning analytics dashboard applications", *American Behavioral Scientist*, Vol. 57 No. 10, pp. 1500-1509.
- Wise, A.F. (2014), "Designing pedagogical interventions to support student use of learning analytics", *Proceedings of the 4th International Conference on Learning Analytics and Knowledge*, March, pp. 203-211.
- Wixom, B. and Watson, H. (2012), "The BI-based organization", in Herschel, R. (Ed.), *Organizational Applications of Business Intelligence Management: Emerging Trends*, IGI Global, Hershey, pp. 193-208.

#### Further reading

- Ferguson, R. (2012), "Learning analytics: drivers, developments and challenges", *International Journal of Technology Enhanced Learning*, Vol. 4 Nos 5-6, pp. 304-317.
- Kitchin, R. (2014), "Big data, new epistemologies and paradigm shifts", *Big Data & Society*, Vol. 1 No. 1, pp. 1-12.
- Roberts, L.D., Howell, J.A., Seaman, K. and Gibson, D.C. (2016), "Student attitudes toward learning analytics in higher education: 'the fitbit version of the learning world'", *Frontiers in Psychology*, Vol. 7, available at: <https://doi.org/10.3389/fpsyg.2016.01959>

#### Corresponding author

Tom Broos can be contacted at: [tom.broos@kuleuven.be](mailto:tom.broos@kuleuven.be)