Making sense of learning data at scale

Book Section

How to cite:


© 2021 Rhona Sharpe, Sue Bennett and Tunde Varga-Atkins

Version: Accepted Manuscript

Link(s) to article on publisher’s website:

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.
Making sense of learning data at scale
Bart Rienties¹, Christothea Herodotou²

¹ The Open University, (United Kingdom), ORCID id 0000-0003-3749-9629
² The Open University, (United Kingdom), ORCID id 0000-0003-0980-1632

Abstract
In many higher educational institutions teachers and managers have unprecedented levels of access to data of learners and their learning. In particular during COVID-19 teachers suddenly had to make sense of what their learners were doing away from the classroom. In this chapter we will provide a range of evidence-based guidelines and best-practices based upon implementing online learning and making sense of learning data on a large-scale institutional level. The Open University UK has been implementing learning analytics at scale and has made learning analytics dashboards available to thousands of teachers. How to make sense of these interactive data streams is complex at the best of times, especially in a situation when most teachers are working remotely. We will reflect on our lived experiences of helping teachers to make sense of learning data, and how their voices are essential for senior managers to build powerful data practices.

Keywords: learning analytics, sense making, data analysis, teachers, management.

Introduction
Across the globe, data is becoming the key currency in business, government, and the wider public sector. Whomever has access to and controls data seems to become the next king or queen in any sector, from on demand streaming services like Netflix to telemetry in cars like Tesla, from technology giants like Apple and Microsoft to online shopping like Amazon. Also, governments are increasingly relying and focusing on big data. During the COVID-19 pandemic it has become evidently clear that all governments at local, regional, national and international levels needed to have good, reliable, and accurate data about and for their citizens. While some higher education institutions (HEIs) and parts of units within HEIs have used data to make sense of what may be happening with their students, in comparison to most contemporary companies and the wider public sector, the use of data is mostly retrospective. For example, at a panel discussion at IMS Global, Phil Richards (2020), Chief Technical Officer from JISC, indicated that:

“[l]ots of assessment activity takes place during the module, formative assessments, summative assessments, and it is no surprise to learn that one of the best predictors of performance in the end of a module assessment is performance at mid-point because they are assessing similar things. I believe there is plenty of evidence that shows that, and the Open University is the exception that proves the rule, they do this extremely well. But in a lot of universities, a lot of mid-module data just lives in a spreadsheet, in the administrative, in the faculty office, until right at the end of the semester when it is too late. So any benefit that could have been got from noticing that student who wasn’t doing so well is completely lost”.

Indeed, until recently most HEIs mostly collated two types of data, often in an isolated manner (Phillips, 2013; Viberg et al., 2018). One type of data, academic learner data, mainly consists
of relatively static data about registered learners, like demographics, prior education, progress and completion data of courses and degrees (Tempelaar et al., 2015). Most of these data are primarily collated and managed to award grades, certificates, and degrees and are maintained for formal quality assurance and quality enhancement processes. While teachers are those who in part generate some of these data, such as class lists of whom is enrolled in a module, as well as sending marking data to the Faculty office, in practice these academic learner data are mainly used by ‘non-teachers’, such as administrators, managers, and policy makers within institutions.

The second type of data, learning data about what learners are (not) doing in a module, is increasingly becoming available to teachers (Herodotou et al., 2019; Rienties, Giesbers, et al., 2016). In particular with the rise of virtual learning environments (VLEs) most teachers in the last 10-20 years have gained more access to affective, behavioural, and cognitive learning data of their learners. For example, with the introduction of VLEs many teachers started to share their lecture slides and videos of their learning activities, create interactive assignments, or initiate discussions on complex problems using discussion forums or blogs (Beetham & Sharpe, 2020; Rienties, Giesbers, et al., 2016). These VLEs often provide standardised metrics of how a learner has, for example, engaged within the VLE, when the last log-in was, and how long a learner has spent on the VLE for a set period of time (such as during the last week). While ‘non-teachers’ could gain access to these learning data, often it is mainly the teacher who could access and use these data to inform their teaching practices. However, as indicated by a range of studies, until recently most teachers do not actively use these data (Herodotou, Rienties, et al., 2020; Kaliisa et al., 2021; Rienties, Boroowa, et al., 2016). This could be explained by multiple reasons, such as teachers may think that there is no need to use them, the data are too complex to understand, or the data stored in these dashboards are not necessarily relevant.

With COVID-19 and the rapid shift to blended and online provision, the amount and quality of data that can potentially be available to teachers and managers in HE about their learners has exploded to unprecedented levels. This raises a number of questions: How can teachers and managers make sense of all these data? What is useful to know about learners and their learning, and what may not be as useful and should be omitted? How should managers help to support and nurture a data culture in their organisation? In this chapter we aim to share some lessons learned from a large-scale implementation of data use by teachers and managers at the Open University UK (OU), the largest distance learning provider in Europe.

How to make sense of large amounts of data?

In the last ten years, an incredible body of research has become available under the umbrella term of learning analytics, which aims to apply the outcomes of analysing data from learners and affect their learning behaviour. A range of systematic literature reviews have found that learning analytics can help to identify which learners are potentially at risk (Bodily & Verbert, 2017), who might need more or less support (Ilenthaler & Yau, 2020), which interventions might be effective for which learner (Ferguson & Clow, 2017; Knobbout & van der Stappen, 2020; Viberg et al., 2018), and how to identify which part of a learning design may not be effective (Mangaroska & Giannakos, 2019). While many of these reviews provide strong evidence and support of the power of learning analytics, most of these applications were conducted in a single-module or small-scale settings (Ferguson & Clow, 2017; Viberg et al., 2018).

In contrast, the research on learning analytics at the OU is based upon our practical experiences of working with large numbers of teachers and managers throughout all 400+ modules and qualifications available within the OU. For example, in one of the first large-scale projects to work with teachers and learning analytics, the Analytics4Action project (Rienties, Boroowa, et al., 2016, p. 1), we argued five years ago “that one of the largest challenges for
learning analytics research and practice still lies ahead of us, and that one substantial and immediate challenge is how to put the power of learning analytics into the hands of teachers and administrators”. While both in the wider literature and our own practical experiences working with thousands of teachers at the OU do indicate some progress, we argue that this challenge is still present, and perhaps becoming even more urgent. In this book chapter we will use the Analytics4Action project as the leading case-study to highlight how teachers and managers might want to consider making sense of rich learning data.

Case study: The Analytics4Action project

In the Analytics4Action project (2014-2016), we realised from the beginning that developing and implementing any dashboard for teachers and managers with learning and learner data, let alone a state-of-the-art learning analytics dashboard, requires involvement of a wide range of stakeholders. As indicated by Rienties, Boroowa, et al. (2016, p. 4), in Analytics4Action “the first step is to bring together the key stakeholders in the module, such as teachers, learning analysts and administrators for the purpose of presenting, unpacking and understanding learning data available taken from various VLE and related systems. This is termed a data touch point meeting [now labelled data support meetings] and the project held four of these with each module over a one-year period”.

As indicated by the Analytics4Action framework (Rienties, Boroowa, et al., 2016), there are six potential steps that teachers and managers of institutions might want to go through to make sense of data about learners and their learning. We will describe each of the six steps below, and note that we are not expecting all institutions to go through all six steps of Analytics4Action, or in any specific order. For some teachers and managers, just starting with the first step might already lead to substantial new insights. For others, it might be already a common practice to look deeply at data, but perhaps there is no evidence hub yet to store and collate these insights.

Step 1: Key metrics and drill downs: what is known about your students?

In the first step teachers and managers start to explore whatever trends are available about their learners in institutional and local data sets and dashboards. By looking at available data and trends important initial insights can be obtained about patterns of behaviour and cognition, which based upon teachers’ experience and intuition might confirm that a particular learning activity, design, course or qualification is working (or not). For example, at the start of the Analytics4Action project in 2014 we showed a range of data options (both dynamic and static data) to our stakeholders about the learners’ progression and usage of specific VLE tools. Gradually and over time, specific tools and options were developed and implemented in the first iteration of the learning analytics dashboard. Rather than showing to stakeholders what is all possible (as the OU collects a lot of data on learners and learning), we focused our initial development of the dashboards based upon what was first and foremost relevant to teachers. Initially, these dashboards just provided key metrics and drill-downs of main data that teachers considered to be relevant, while over time these led to more advanced learning analytics systems like OU Analyse, forecasting future learner performance (Hlosta et al., 2020), as explained in detail in Box 1.

For example, based upon the feedback from stakeholders, our early version of our Analytics4Action dashboard allowed teachers to analyse how many unique learners still logged into the VLE in a respective week, relative to other modules and previous implementations of the respective module. During three data support meetings, teachers worked with other stakeholders and data experts through a range of dynamic and static dashboards. For example, a teacher looking at data in say week 6 might conclude that, relative to the previous year, a lower percentage of learners logged into the VLE. This could be an expected outcome if, for example, a teacher has made a learning design change in the module in week 5-7 (e.g., study break, an offline activity). It could also be an unexpected outcome if no changes had been made to the learning design in week 5-8, or perhaps a new
learning design change (e.g., introduction of a discussion task) did not have the expected impact. Only the teacher would be able to ask critical questions and make sense of patterns within the data. Based upon a dialogue with the teachers, data experts developed follow-up drilldowns and metrics to allow teachers to further unpack why a particular trend in the data was (not) happening in (near) real-time. By iteratively fine-tuning the dashboards together with teachers, a co-constructed approach was adopted to make sense of the emerging data.

### Box 1 Application of advanced learning analytics with OU Analyse

While in previous versions of our Analytics4Action dashboards teachers had only rudimentary access to some data (e.g., how many learners have been active in a particular week), in OU Analyse – an advanced learning analytics dashboard developed by the OU – teachers and managers have access to fine-grained and detailed predictive learning analytics data on a weekly basis for each learner/module. For example, Figure 1 presents a large-scale introductory module on social sciences of 2,762 learners that was implemented just a month before the UK went into lockdown in March 2020.

Based upon these data imagine the following scenario: You are a teacher or manager of this module responsible for a sub-set of 20 learners. Furthermore, the bar charts indicate the average assessment scores of the various module assessments relative to the previous cohort. Until week 4 (1st week of March 2020), your learners were substantially more active in terms of average numbers of clicks relative to the previous cohort. So, is there anything to worry about?

From week 4-8 onwards, your learners were less active relative to the previous cohort, perhaps due to the start of lockdown. In the remainder of the course, there are not a lot of differences on an aggregate level between the current and previous cohorts. Yet, at the end of the module, there were some substantial differences in learner outcomes, whereby 1,156 (41% vs 46% last year) of learners passed, 618 (22% vs 28%) failed the course, and 1,086 (39% vs 47%) withdrew from the course. These pass-rates must be seen in light with the open access policy of the OU, where anyone without any qualification or degree can register for introductory courses like this one.

Teachers also have access to the predictions of their learners through the OU Analyse dashboard that predicts whether a learner is going to submit (and pass) the next assignment, using a traffic-light like visualisation (not illustrated). For example, one learner called Agnes did not submit the first assignment (in OU jargon this is called Teacher Marked Assignment, TMA). Over time more learners were flagged as at risk, whereby for example in week 16 four out of ten learners were expected to submit the third assignment, and six would not. In OU Analyse, teachers and managers have access to both traffic lights and simple up/same/down/flagged icons highlighting the trend over time for each learner. A teacher can subsequently zoom into this respective learner to obtain a detailed breakdown of the underlying behavioural and cognitive patterns of a learner, and why a particular learner might be flagged at risk (not illustrated).

For example, if a teacher would click on learner Agnes it would reveal that relative to the average cohort the engagement of Agnes is more infrequent. She is very active in certain weeks (e.g., week 2, 6, 14, 18, 25) while in other weeks she barely engages online. In part, this can be explained by the learning design decisions made in this course, as alongside a range of interactive online materials, two printed books can be used by learners. It might of course be that Agnes was actively studying using these printed books in the other weeks. One of the smart tricks of OU Analyse is that its combination of machine learning approaches can predict which learning activities are key for successful completion of an assessment, and which ones are not so relevant, as well as how a learner has previously
performed. Based upon thousands of successful and unsuccessful learner paths, OU Analyse gives information to the teacher whether or not key learning activities have been undertaken by Agnes (Herodotou, Rienties, et al., 2020; Hiosta et al., 2020). While early predictions until week 16 indicated uncertainty about whether Agnes would submit, the submission of TMA2 and accessing the homepage activity in week 14 eventually led OU Analyse to assume that she would submit TMA3, which indeed she did with a good assignment score of 80%. She successfully completed the module at the end, even though she was flagged as at risk in week 16. This example shows that learners make complex, non-linear and (sub)conscious decisions when learning online, and systems like OU Analyse can help to inform teachers and managers to support learners.

Step 2: Menu of response actions: what actions can a teacher take in response to data?

The second step aims to provide teachers and managers with a menu with response actions to take pro-active action (where needed) based upon the learning data provided. As with the COVID-19 pandemic, while it is initially difficult to get all the right data in one place to analyse, once the data are in place, these are only relevant if they can lead to actionable insight. For example, in Analytics4Action we conceptualised the menu of response actions based upon the Community of Inquiry framework by Garrison and Arbaugh (2007) and Cleveland-Innes and Campbell (2012). Some of the responses from these data support meetings could be positioned as a teacher presence (e.g., running an additional online session to clarify key concept X), others as a cognitive presence (e.g., altering the description and narrative of a key learning object Y), while others could be to enhance the social presence between learners (e.g., running an informal webinar party). Obviously, some of these response actions might be more costly or complex to implement relative to others. Furthermore, for some of these interventions there is already some evidence that these kinds of interventions make sense, while for others the effects are still unknown. Note that due to the unique nature of the OU, these interventions must be done at scale, and therefore each intervention is costed and subsequently approved depending on the nature and type of intervention. Individual teachers at your institution may have much more freedom and opportunity to start with an intervention as they see fit.

Step 3: Menu of protocol: how to evaluate the impact of your intervention?

After a teacher has selected a particular learning design or active-intervention strategy based upon the data analysis, the third step in Analytics4Action is to determine the research protocol to unpack and evaluate the impact of these intervention strategies. As illustrated with the examples above, an institution might decide to intervene on an individual learner level (Herodotou, Rienties, et al., 2020), across a particular group of learners (Herodotou, Naydenova, et al., 2020), or even across all learners with a learning design intervention (Rienties, Boroowa, et al., 2016; Rienties et al., 2018). Initially, when we implemented the Analytics4Action framework we expected that over time the OU would increasingly start to implement more evidence-based approaches like randomised Control Trials and A/B testing. While some progress has been made in a small number of interventions (e.g., Herodotou et al., 2017; Herodotou, Naydenova, et al., 2020), overall most of the interventions conducted to date could be classified as quasi-experimental (compare cohort 2 to the previous cohort 1) or apply to all. This obviously has the advantage to a teacher that the intervention protocol is simple, and perhaps fair as all learners get the same treatment. At the same time, from an evaluation protocol standpoint, it is difficult to make inferences as to whether the intervention worked (or not), and for whom.
Step 4: Outcome analysis and evaluation: how do you know it worked?

Obviously a key question in any intervention is to explore the impact of an intervention, or put simply ‘did it work?’. At the OU teachers often use a combination of scholarly research with critical reflection on lessons-learned using the yearly quality enhancement monitoring to determine whether (or not) a particular intervention strategy has worked or not (Rienties, Cross, et al., 2016). For example, within the Analytics4Action approach teachers frequently reflect together with the learning design team whether or not a particular intervention has worked (Hidalgo & Evans, 2020). For example, in 2017/18 the learning design team worked with 49 modules, running 136 module data support meetings for a total of 128 teaching staff. As illustrated in the seven selected interventions by Hidalgo and Evans (2020), several showed a positive effect relative to the previous cohort. One way to further strengthen the outcome analysis is to triangulate the quantitative data with the lived experiences of learners. A lot of scholarly projects at the OU aim to obtain a deeper understanding of why some interventions might work for some groups of learners, but perhaps not for others.

Step 5: Evidence hub: how to share experiences with other teachers?

Another way to gather evidence of whether an intervention has worked or not is to combine the insights from various interventions done within a unit, or even across the institution. Therefore, Analytics4Action proposed in the fifth step to collate the various experiences from teachers in a so-called evidence hub (De Liddo et al., 2012; Ferguson et al., 2016). This step probably will need to be initiated by senior management to ensure that there is both an infrastructure and an incentive for teachers to share these practices. Even if an intervention, for example running an informal webinar party to encourage social presence might work in four modules in sociology and physics, there is no guarantee that such an intervention would also work in say a business module. By collating the various intervention strategies, approaches and lessons-learned teachers can make a more informed decision when they identify a particular issue from the Analytics4Action data support meetings. In the OU, teachers are encouraged to write up their experiences in an online repository called Scholarship Exchange.

Step 6: Deep dive and strategic insight: are the findings systemic or specific for one module?

A sixth and final step that probably is more on a strategic level rather than on an individual teacher level is to conduct a deep dive analysis based upon the specific (though perhaps small) interventions made by teachers. For example, within the OU a range of big data studies (Nguyen et al., 2017; Rienties & Toetenel, 2016) were conducted to determine how learning design decisions made by teachers influenced learners engagement, behaviour and cognition. These studies indicated that learning design decisions made by teachers had a fundamental and strong impact on both the behaviour of students and their learning outcomes. In particular, 69% of weekly behaviours by learners were predicted by how a particular teacher designed a module. When teachers designed communicative learning activities (i.e., learner to learner, teacher to learner, learner to teacher) this had a strong predictive impact on both engagement and learning outcomes. Similarly, in a study using logistic regression on 123,916 undergraduate learners in 205 modules over several semesters from 2015-2017 (Nguyen et al., 2018) showed that including a study-break had a positive impact on retention, while including assessment preparation and revision weeks at the end of a module had no impact on retention (in contrast to a popular belief in the OU). These kinds of big data and deep dive studies need substantial support from senior management as well as substantial data modelling experience, but can provide crucial insights about whether a common educational practice embedded in an institution is actually working or not.

Conclusions

As highlighted in a recent study about how people make sense of data, people are more inclined to listen to other people who have experienced a particular problem or issue, rather
than relying on facts and data alone (Kubin et al., 2021). By having teachers and other stakeholders working together, and sharing these experiences with colleagues within their own department and network, they can over time convince others of the usefulness, affordances and limitations of some of these approaches and tools (Herodotou, Rienties, et al., 2020; Rienties et al., 2018).

The initial Analytics4Action project became business as usual after positive reactions by staff. Since 2016 Analytics4Action programme led by the OU Learning Design team has supported over 213 modules, 603 staff and 611 data support meetings in total, impacting learning designs and interventions for around 158,000 students (Hidalgo, 2021; Hidalgo & Evans, 2020). Furthermore, the sharing of learning analytics practices from the OU with other HEIs has resulted in an impact on the understanding, learning and practice of 1541 university educators over a dozen countries, including Belarus (Olney et al., 2020), China (Olney et al., 2021), Kenya (Mittelmeier et al., 2018), and South Africa (Greyling et al., 2020).

Given the value data dashboards can bring to teaching, we would encourage managers to promote teachers to start to engage with data by using a common learning tool. In most tools there is information as to how each user has been using a particular function, such as the average engagement time, or clicks on a particular day. This information can be an indication of how engaged learners are and what information is exchanged with others. Accessing these insights can be a first step towards experiencing how data can inform the teaching practice and facilitate an understanding of what students may be doing when studying from a distance.

Perhaps more importantly, it will raise a lot of questions in terms of what other data might be important to understand the learner journey, and what data infrastructure(s) and expertise might be needed to build capacity within an organisation to make sense of these data. Learning analytics is just one source of information that should be used alongside other insights an institution has access to, such as students’ previous performance, attendance of webinars, submission of assignments or direct phone or email communication, to ascertain whether a student is progressing well or whether additional support should be provided.

**Future directions**

It is acknowledged that the process of engaging with learning analytics is not straightforward and as illustrated in our case-study of Analytics4Action may require substantial support from the management, IT and educational experts for teachers to understand the significance of the data and act upon data. For example, redesigning their teaching activities or providing personalised support to students in need may be difficult to achieve by a one individual teacher, but by working together within an organisation this can be achieved over time. We would encourage educational institutions to invest resources in gathering or restructuring available student data in ways that enable stakeholders to access these and use data to inform their practices. For example, a group of teachers could act as ‘champions’ to familiarise themselves with the data, liaise with educational and IT experts and develop an in-depth understanding of what data means and how it could be used. They could then act as the agents communicating this knowledge and expertise to their peers and helping them access, understand and use data in their practices. What should be clear is that engaging with student data in a systematic manner can empower staff, especially when they are teaching from a distance or blended, by enabling them to monitor and keep track of what their students are doing, how they are performing and what difficulties they may be facing. Data can inform any decisions educators make and justify the choice and application of certain interventions to support students from a distance.

**References**


Olney, T., Endean, M., & Banks, D. (2020). Evaluating the impact of the learning design and course creation (LDCC) workshop on the participants of the enhancement of lifelong learning in Belarus (BEL) project. Proceedings of the final conference on the Erasmus+ project,


Figure 1 OU Analyse data from a social science module during the pandemic