

# Chapter 6

## Overcoming the MOOC Data Deluge with Learning Analytic Dashboards

Lorenzo Vigentini, Andrew Clayphan, Xia Zhang and Mahsa Chitsaz

**Abstract** With the proliferation of MOOCs and the large amount of data collected, a lot of questions have been asked about their value and effectiveness. One of the key issues emerging is the difficulty in the sense—making from the data available. The use of analytic dashboards has been suggested to provide quick insights and distil the large volume of learner interaction data generated. These dashboards hold the promise of providing a contextualized view of data and facilitating useful research exploration. However, little has been done in defining how these dashboards should be created, often resulting in a proliferation of systems for each new research agenda. We present our experience of building MOOC dashboards for two different platforms, namely *Coursera* and *FutureLearn*, motivated by a set of design goals with input from a diverse set of stakeholders. We demonstrate the features of the system and how it has served to make data accessible and useable. We report on problems faced, drawing on analyses of think-aloud sessions conducted with real educators, which have informed our dashboard process.

**Keywords** Learning and teaching data · Dashboards · MOOCs · Visualization · Sense-making

---

L. Vigentini (✉)

School of Education, UNSW Australia, Kensington, NSW, Australia  
e-mail: l.vigentini@unsw.edu.au

L. Vigentini · A. Clayphan · X. Zhang · M. Chitsaz  
Portfolio of the Pro-Vice Chancellor Education, UNSW Australia,  
Kensington, NSW, Australia  
e-mail: a.clayphan@unsw.edu.au

X. Zhang  
e-mail: Xialisa.zhang@gmail.com

M. Chitsaz  
e-mail: m.chitsaz@unsw.edu.au

## Abbreviations

HCI	Human computer interaction
HTML	Hypertext markup language
JSON	JavaScript object notation
MOOC	Massive open online course
SUS	System usability scale

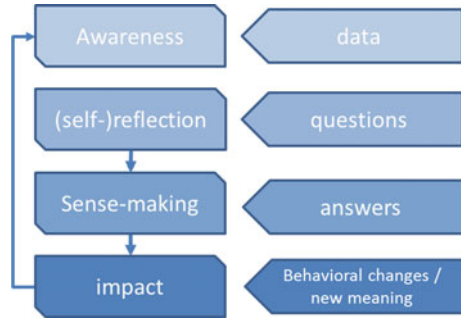
## 6.1 Introduction

Over the past few years, MOOCs (Massive Open Online Courses) have become the center of much attention both from the public media (Cormier 2009; Bates 2012; Cooper and Sahami 2013; Dominique 2015) as well as the research community (Daniel 2012; Amo 2013; Baggaley 2013; Davis et al. 2014; Drachsler and Kalz 2016). Despite the hype of big data in education and the potential associated with the ability to collect and analyse large amounts of information about students' learning behaviours (Arnold and Pistilli 2012; Verbert et al. 2013b), one of the biggest limitations is finding how to expose this data in a meaningful and relevant way for different stakeholders, being students, instructors, researchers or developers (Duval 2011; Dernoncourt et al. 2013; Verbert et al. 2013b). Visualization of data and the ability to manipulate visualizations has been demonstrated to provide useful insights (Pauwels et al. 2009; Duval 2011). These are grouped into dashboards, which have been advocated to enable individuals, researchers and policy makers—quick insight into data. Dashboards are often justified by themes such as providing awareness, reflection and sense making (Allio 2012). However, as dashboards have become popular, they have also brought an over-expectation of benefits with them (Davis et al. 2014; Leon Urrutia et al. 2016). In fact, dashboards in and of themselves do not automatically confer learning or awareness gains. It is a combination of effective design, requirements elicitation, and an understanding of stakeholder objectives, which may permit dashboards to be potentially powerful tools for aiding data exploration.

Learning analytics has been suggested to provide a theoretical framework to make sense of learner's interactions with online courses (Verbert et al. 2013a; Corrin et al. 2015; Corrin and de Barba 2015; Drachsler and Kalz 2016). In fact, one of the purposes of learning analytics is to visualise learner activity so that educators can make informed decisions about possible interventions (Clow 2013; Bayne and Ross 2014; Stephens-Martinez et al. 2014). Verbert and colleagues represented the process aided by learning analytics in Fig. 6.1. In order to achieve impact in the improvement of learning and teaching, data is the key entry level in order to articulate appropriate questions.

In this chapter, the authors report about their experience of developing an analytics dashboard for two different MOOC platforms (Coursera and FutureLearn,

**Fig. 6.1** Adapted from the learning analytics process model (Verbert et al. 2013a)



two leading MOOC providers in the USA and the UK respectively), focusing on the challenges of the process and providing useful insights for others embarking on similar initiatives. Firstly the work is placed in the context of existing work in the field of Learning Analytics where the problem is explored in more detail, discussing key aspects in relation to prior works. Then, the development and implementation process will be described focusing on the process and some of the challenges that the team faced, and finally the results from a small case usability study will be reported to bring support to further development of this work.

## 6.2 How This Work Enhances the LA Field

The sort of dashboard creation process for MOOCs reported here is not new; however it has a number of distinctive features, which provide an opportunity to extend the work carried out in the learning analytics community.

First of all, a number of design goals were established at the outset to cater for multiple stakeholders in a holistic way: this moves away from the specific focus on the teachers (Stephens-Martinez et al. 2014; Corrin et al. 2015), the students (Arnold and Pistilli 2012; Corrin and de Barba 2015; Kia et al.) or institutional research/BI practitioners (Campbell et al. 2007; Mohanty et al. 2013).

Secondly, we had an opportunity to work with multiple platforms which, given the restriction with the data sharing agreements imposed by different providers, makes collaborations between institutions less likely to occur. With exception of a few (Siemens et al. 2011; Cobos et al. 2016), most of the other works published thus far only focus on one platform at the time.

Thirdly, we leveraged on an educational lens to integrate institutional reporting matched with an academic analytics perspective. This allowed focusing on questions at multiple levels of analysis (micro, meso and macro) depending on the relevance to the individual, the course or the institution. This approach is original because the design process focused on generalizable features first, relevant not only in MOOCs, but extending to other learning management systems or learning technology tools as well.

### 6.2.1 *Problem Statement*

When the data team at our institution was tasked to report the findings from our MOOC courses (20+ courses now), it quickly became apparent there were a number of different and competing goals. These ranged from broad institutional questions, down to fine-grained details within particular modes of assessment. As such, the process quickly led to an explosion in the number of questions being raised.

These sorts of problems are common and have been chronicled by others (Seaton et al. 2013, 2014; Verbert et al. 2013b; Qu and Chen 2015; Leon Urrutia et al. 2016) leading to the recommendation of standardisation at various levels from the data (Dernoncourt et al. 2013; Veeramachaneni et al. 2014a) to the sets and types of analytics (Duval 2011; Siemens et al. 2011; Verbert et al. 2013b) or the types of visualisations deemed useful (Stephens-Martinez et al. 2014).

In our context, this rapidly drove a nail into initial ad hoc or exploratory approaches to analysis that were meant to provide quick solutions to some of the questions raised by instructors, which didn't necessarily satisfy other stakeholders, nor make the process efficient.

When each of the courses closed, as expected, the analytical team faced high expectations to produce interesting results from the analysis, particularly given that at the time there was not a fully-fledged dashboard available. Team members faced a number of challenges in order to come up with solutions. For example, course academics (instructors) were keenly waiting on analysis and reports on their course data and their research questions; Educational developers wanted answers about their design; and Senior Managers wanted to know whether their investment had significant returns. In addition, the process to actually get the data had a significant delay, adding pressure to the process and the team with a number of key issues:

- Manual data requests required to obtain the data from the educational platforms (Coursera) and bespoke third party systems after course completion, see Fig. 6.2.
- A steep learning curve due to complexities/volume of data from each system.
- Unsustainable processes: Short timeframes to answer a large number of research questions in ad hoc report formats.
- Concurrent priorities with teaching/research focus: institutional research versus course instructors' own research questions.
- Lack of transparency in requests from stakeholders and interested parties.

The need to establish a methodology to build a reproducible workflow became apparent—going from data extraction, to transformation to delivery. We were confronted by a number of problems, which we used as pointers to define our initial needs analysis of the requirements and scope:

- *Time*: A large amount of effort is required to transform data into a useable format. Doing this by hand is not a long-term solution. How can this be automated? What needs to be automated? What should we strive towards?

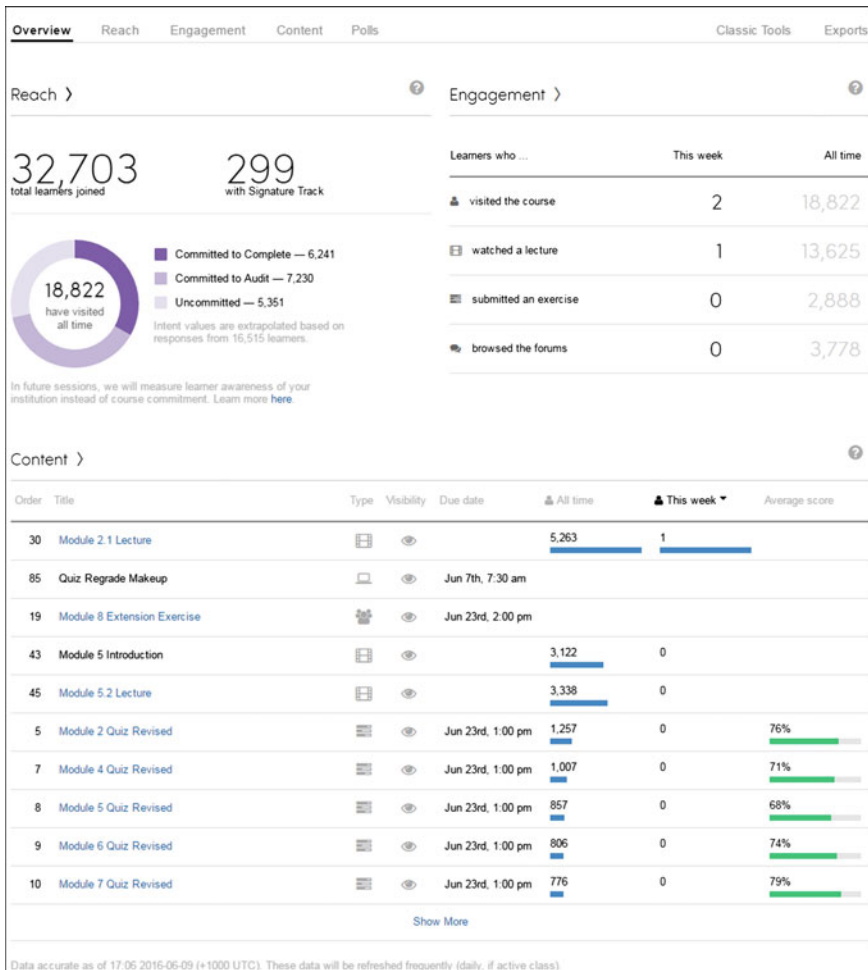


Fig. 6.2 Screen capture of the Coursera analytics dashboard for one of our courses

- *Volume*: There is an ever-growing base of data and variants in the data schemas. What are the common questions being raised, and is this collated somewhere? Are these questions transparent between different members of the data team working on them and the stakeholders asking them?
- *Competencies*: How can we best utilize resources, both in terms of systems and people? Are stakeholders able to understand the output?

In order to address the above challenges, the team developed a process grounded on five building blocks:

1. Articulate a reporting framework, which keeps into consideration the needs of different stakeholders.

2. Develop a semi-automated data transformation workflow.
3. Determine which tool is most appropriate for the needs of stakeholders.
4. Design and build a dashboard, which provides flexibility in the exploration and addresses the majority of the questions raised by stakeholders.
5. Provide a framework for the scalable and sustainable reproduction of the process for each MOOC.

## 6.2.2 *Related Work to Dashboards*

Although analytics dashboards to support organisations in their decision making processes have been around for some time (Pauwels et al. 2009), those focusing on learning analytics have a shorter history. A useful classification according to their purpose was proposed by Verbert et al. (2013b). The first type of dashboard supports traditional face-to-face lectures; another type of dashboard supports face-to-face group work; further, another type of dashboard is used to support awareness, reflection, sense-making and behavioural changes in online or blended learning. In this chapter we focus specifically on the latter. In fact, in the current MOOC space there are relatively few examples of learning analytics dashboards, with mostly disjointed approaches and ad hoc implementations. For example, Coursera provides a dashboard to instructors with a live view of the data, but the granularity of the information does not necessarily cater for all stakeholders needs or wants (such as the ability for instructor-specific stratification—see Figs. 6.2 and 6.3–*left* for examples of high-level views of course data). Other vendors, such as FutureLearn, take an even further removed option, with no visual dashboard, but rather just a list of key course metrics (Fig. 6.3–*right*).

Although these representations provide a high-level view of MOOCs that may be suitable to get a general sense and may satisfy academic managers; educators are likely to ask more probing and sophisticated questions about contributing factors leading to certain patterns of engagement (Stephens-Martinez et al. 2014).

On the other hand, other platforms like EdX (Seaton et al. 2013; Ruiz et al. 2014; Fredericks et al. 2016; Pijeira Díaz et al. 2016) recently added analytical plug-in modules, that users can install on their own systems, to provide detailed views of how learners engage with the platform. These however may be too detailed and/or complex to cater for the casual user.

There are also examples of external visualizations tools and dashboards (Veeramachaneni et al. 2014b; Qu and Chen 2015; Cobos et al. 2016; Davis et al. 2014; Leon-Urrutia et al. 2016; Kia et al.), which provide specific representation of behaviours in MOOCs, however the majority of these tools are not open. This often means that teams in different institutions end up replicating similar processes. Furthermore, given the specific reasons why the dashboard is developed, it is natural to ask whether the information represented conveys useful insights to stakeholders outside initial developments.

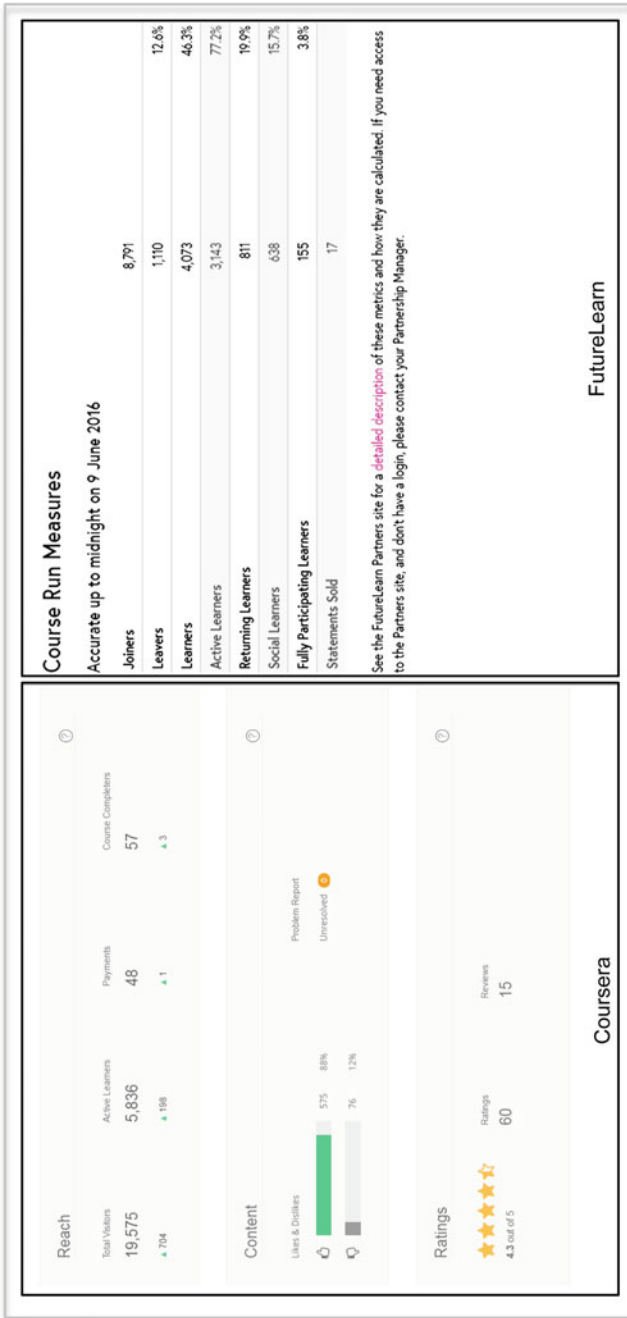


Fig. 6.3 On the left, key figures from the Coursera dashboard. On the right key metrics from FutureLearn

**Table 6.1** Data sources/data tracking in learning analytics from various published works

Data sources	Data tracking
– Artefacts produced <sup>a,b</sup>	– Application logs <sup>a</sup>
– Social interaction (forums) <sup>a,b</sup>	– Camera <sup>a</sup>
– Resource use <sup>a,b</sup>	– Microphone <sup>a</sup>
– Time spent <sup>a,b</sup>	– Depth sensors <sup>a</sup>
– Exercise/Test results <sup>a,b</sup>	– Biofeedback <sup>b,c</sup>
– Assessment/Course grades <sup>b</sup>	– Eye trackers <sup>d</sup>

Sources <sup>a</sup>(Verbert et al. 2013a); <sup>b</sup>(Romero et al. 2008); <sup>c</sup>(Govaerts et al. 2012); <sup>d</sup>(Vatrapu et al. 2013)

Nevertheless, as noted in Verbert and colleagues’ comparison of different learning analytics dashboards (Romero et al. 2008; Govaerts et al. 2012; Vatrapu et al. 2013; Verbert et al. 2013a) there are common elements that can be used as a starting point. Table 6.1 provides a summary (albeit not comprehensive) of different types of sources and types of tracking from currently published literature.

### 6.3 Design and Development of the Dashboard

A number of design goals were established from the process of exploring the problems around the analytics space. These were: (1) the need for reproducible processes; (2) flexibility; (3) transparency and (4) extensibility. Based on these, we proposed a framework to enable data exploration efforts to take place that attempted to be platform agnostic. A number of questions emerged which are explored in more detail in the next sections.

Before stepping into the description of the rationale and method for planning, designing and developing the dashboard, it is necessary to quickly describe the type of data available in MOOCs. Table 6.2 provides an overview of four categories of data available in MOOCs: these are commonly valid across MOOCs and Virtual Learning platforms, but the granularity of the details vary. For example while rich demographic data may be available to institutions for credit-bearing courses that they offer, in the MOOC space, information about participants is limited and in general rather sparse. On the contrary, the logs of online activity in MOOCs are more sophisticated than most on-campus blended courses. Notably, as demonstrated in Table 6.1, a large proportion of works published, focused primarily on activity data and performance data, which provide a limited window on the learning experience (Vigentini and Zhao 2016). Coursera and EdX provide a great deal of information about the interaction with videos, however FutureLearn at this point in time does not provide any details about ‘in-video’ behaviors.



**Table 6.2** Types of data available in MOOCs and other learning platforms

Demographic data: general student demographics, including age, gender, language, education level, and location. Demographic data is commonly acquired during the registration process, and additional demographic data can be acquired via feedback surveys	Performance data: student performance based on both formative and summative assessments. This is collected from homework, quizzes, and exams, but it also includes results from pre-course surveys/activities designed to diagnose student knowledge before they take the course
Activity data: how and when students are using the resources, such as watching videos, reading material, submitting homework, taking quizzes, or using the discussion forum. Most platforms break down usage by content and media type (i.e. page views, assignment views, textbook views, video views)	Feedback data: student feedback coming from polls, surveys and comments in forums. Additionally information about student learning goals and motivation and intended use of the material

### 6.3.1 *What Is the Most Appropriate Tool to Build the Dashboard in?*

A preliminary analysis into a number of commercial visualization platforms was conducted. This consisted of assessing (1) availability, (2) flexibility, (3) adaptability, (4) export features and (5) data privacy. The following tools were examined: Ubiq; Tableau; SiSENSE; The Dash; Dashzen; Ducksboard; Klipfolio; Leftronic; Qlik; Drillable; Logi and Infocaptor.<sup>1</sup>

We also examined the issue of data privacy. A detailed analysis made clear that not one tool could solve all the issues we had, with most tools coming short with some features or a hefty price tag.

A systematic evaluation and review of the comparison is beyond the scope of this chapter; however it should be noted that different products had fundamental differences in the visualisation capabilities, management of the analytic workflow and varying technical affordances. Therefore we opted to build a lightweight scaffold that allowed multiple tools to be used. This web-based architecture was driven by a theoretical representation of the data as well as the intent to produce a range of visualisation tools (widgets) that helped to explore the data from multiple perspectives. This process was strongly informed by the general framework in Siemens et al. (2011) and the attempt to integrate visualisations at Harvard, Stanford and Berkeley (Dernoncourt et al. 2013; Pardos and Kao 2015).

For what concerns the data format, we considered the MOOCDB data structure (Veeramachaneni et al. 2014a), which seemed to promise a unified approach for data coming from EdX and Coursera, but had to discard the options for two main reasons: (1) the pace at which Coursera changed the format of their data exports made the process unwieldy and difficult to maintain the code up to date; and (2) the

<sup>1</sup>Please refer to the note at the end of the chapter with links to the respective websites.

level of transformation required would have led to a substantial loss of information even before we began the analysis of the data. Furthermore, the standard proposed did not actually account for considerable platform differences nor the contextual assumptions determined by the design of the user interface.

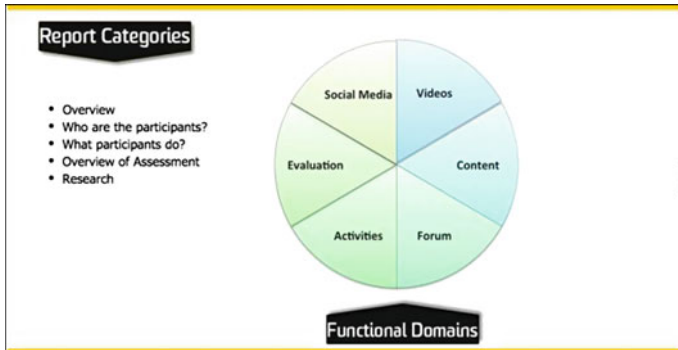
### 6.3.2 How Can We Best Answer the Questions Posed?

After considering the list of questions asked and informed by the data available we developed a framework to allow different ways to organise, display and get to the required information. The framework is based on two top level groups: report categories and functional domains. These are shown in Table 6.3 and then visually in Fig. 6.4 providing the basis for component re-use so that visualizations built for one course, could be easily adapted and re-used in another course. Figures 6.5 and 6.6 provide examples of the form of visualizations we have built to date.

*Report categories* represent standard reports organised by specific labels. In general the categories refer to the course as a whole. The representations under this label focus on questions that instructors or academic managers would want answers to in a broad way.

**Table 6.3** Broad reporting framework based on categories and functional domains

Report categories	Functional domains
<p>‘<i>Overview</i>’ comprises the funnel of participants to model the student’s journey in the course, from when a learner becomes aware of the course to the completion of it. It covers the size of the MOOC, breakdown of the participants by status and achievement level, type and level of activities</p> <p>‘<i>Who are the participants</i>’ covers visualisations related to user characteristics: stratification on demographic variables such as age, gender distribution, and education and employment status</p> <p>‘<i>What participants do</i>’ involves user activities and use of course content: videos, forums, quizzes and peer assessments</p> <p>‘<i>Overview of assessment</i>’ looks at the cognitive aspects of student performance including both formative and summative assessments</p> <p>‘<i>Research</i>’ provides insights drawn from data mining techniques. For example, this includes cluster analysis based on users’ activities</p>	<p>‘<i>Videos</i>’ focuses on the daily, weekly, monthly use, broken into actions of downloads and views. Videos are listed by module/week and can be drilled down to individual levels.</p> <p>‘<i>Content</i>’ explores the daily, weekly, monthly use of course aspects outside of videos.</p> <p>‘<i>Forums</i>’ contain daily, weekly, monthly use broken into actions of posts and comments. It also provides insight into participant interactions</p> <p>‘<i>Activities</i>’ focuses on the engagement opportunities, a student has in the course to help them learn. It refers to both formative and summative activities</p> <p>‘<i>Evaluation</i>’ shows the methods used in the course to evaluate the student experience and effectiveness of the MOOC</p> <p>‘<i>Social Media</i>’ contains detail about the reach of the MOOC through certain distribution channels</p>



**Fig. 6.4** Dashboard frontend for our Coursera courses showing the 3 main entry points for visualizations: report categories, functional domains and the full site map from the right panel activated by pulley button (*arrow*)

*Functional domains* arrange reports and visualisations according to what their purpose is in the MOOC. The key difference of analysis and representations under this label is the level of granularity of the analysis. Instructors with specific questions about their content or educational developers/learning technologists detailed questions about what worked or not in their learning design. Furthermore, we designed the dashboard as a data distribution point (following in the footsteps of moocRP (Pardos and Kao 2015) allowing individuals wanting to explore further the data to download the data from the visuals provided.

However, unlike moocRP we felt that the workflow to make data available needed to be less sophisticated allowing individuals to extract what they needed after they identified valuable visualisations, so that they could quickly drill-down to answer their questions.

### 6.3.3 How Can We Make Use of Effective Visual Design?

We largely based our framework and dashboard design on the principles from the work by (Duval 2011; Siemens et al. 2011). In summary, these were: (1) provide viewers with the information they need quickly and clearly, (2) stay away from clichés or gimmicks; (3) focus on what is important; and (4) align to educational objectives and learning goals.

The technical solution took into account fundamental differences in the data provided by the two platforms and also two different workflows.

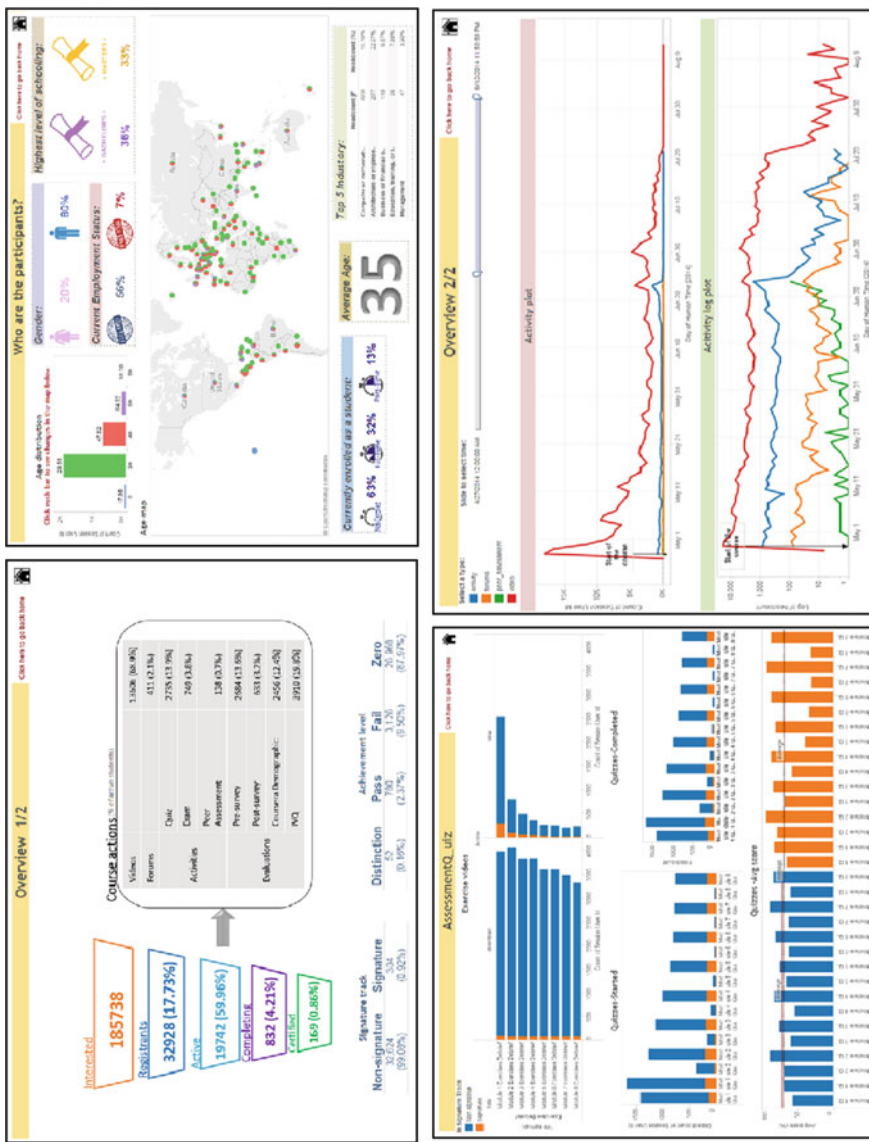


Fig. 6.5 Examples of visualizations from the MOOC Dashboard. *Top*—overview and demographics; *bottom*—assessment performance and timeline of activities

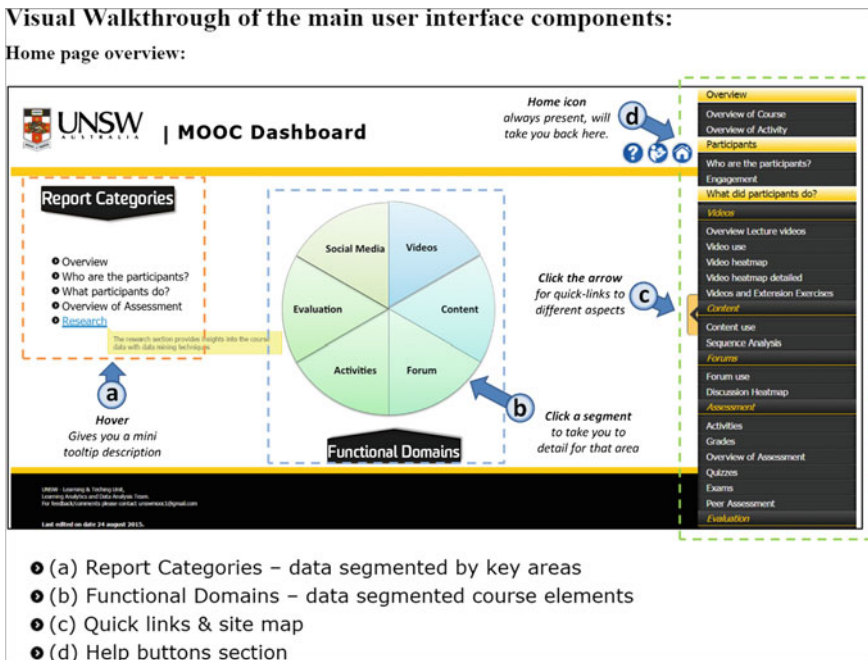


Fig. 6.6 Visual walkthrough of the dashboard developed for Coursera

## 6.4 Building the Dashboards

The development of the building blocks (or widgets) and the aggregation into dashboard panels was an iterative process based on four stages:

1. Design the dashboard elements/blocks
2. Prototype the visualizations
3. Test the dashboard with different users to collect feedback
4. Evaluate and iterate over the design.

The dashboard framework for Coursera is relatively simple. Two HTML pages (a home page and a panel page) are the frontend organizing the content and menus. In the backend there is a single JSON configuration file per course that is used to dynamically populate the html elements and the critical component is a set of visualizations created and packaged in Tableau, which are served based on the configuration file. Examples are shown on the next page (Figs. 6.6 and 6.7).

The dashboard for FutureLearn is still at the early stages of development, but it uses a different architecture, relying on Shiny dashboard and R scripts to generate the various visualizations (Fig. 6.8). These are then framed as widgets and organized according to a similar taxonomy as the Coursera dashboard.

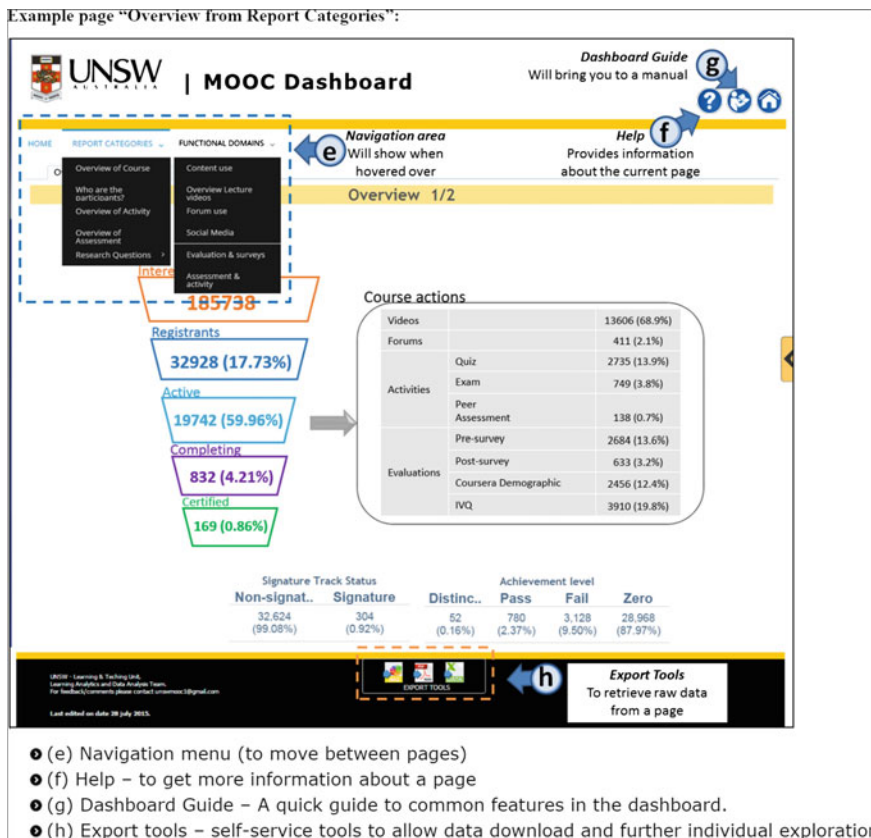


Fig. 6.7 Visual walkthrough of the dashboard developed for Coursera

## 6.5 Evaluating the User Experience

Initial feedback from instructors has been very positive, however we have also been evaluating the user experience more formally by designing an in-house process for a think-aloud and cognitive walkthrough protocol (Wharton et al. 1992; Fonteyn et al. 1993; Rieman et al. 1995; Azevedo et al. 2013). Think aloud methods are cost-effective, robust, flexible and relatively easy to administer (Nielsen 1994; Conrad et al. 1999).

Based on the review of the literature and the questions raised by our instructors it was evident that instructors need to understand the effectiveness of resources, activities, grading rubrics and support methods in relation to the set learning outcomes and in order to continuously improve the course (Churchill et al. 2013). Academic managers are more interested in the ‘bigger picture’ and draw comparison between credit-bearing courses offered by the university and the courses offered for free as MOOCs. Another dimension of interest for our courses was the



**Fig. 6.8** Snapshot of two widgets for the FutureLearn dashboard representing the quartile distribution of quiz answers and the prevalence of high degrees in participants

**Table 6.4** Overview of activities, in terms of the number of participants who answered the question successfully, the average time taken to do the activity, and the most common pathway utilized to obtain the answer reported

Activity	Correct response ratio	Mean time	Most common method to arrive at response
(1) How many times on average was the first video from the first week/module viewed by non-signature track students?	7/11	4:45	Functional domains
(2) What percentage of people passed the course?	10/11	2:43	Report categories
(3) How many students finished the peer assessment for <exercise x>?	11/11	2:03	Report categories
(4) How many people made a forum post on <date x>?	9/11	3:42	Functional domains

evaluation of effectiveness for different pedagogical approaches, which led to the design of a set of test activities to determine whether the dashboard fulfilled its intended purpose.

The test activities were designed on real scenarios presented to us by educators and project support staff, yielding an element of authenticity. The four questions



(see Table 6.4) were intended to be simple, but non-trivial as they required some level of exploration and integration of information.

### 6.5.1 Study Design

To gain feedback about the dashboard interface, we developed a protocol to inform a quasi-experimental interaction study bringing in real-users to test the dashboard, and drew on human-computer interaction (HCI) techniques—a think-aloud process combined with a SUS usability questionnaire (Brooke 1996). Each session was conducted individually, accompanied with screen and audio capture (using QuickTime recording on the computer utilized for testing).

The protocols aimed to standardize the testing sessions and allow for some comparability in the observations. The procedure began with a background questionnaire, to collect general information about the participants' computer skills and past dashboard experience. Following this, the participants were each given a sheet that explained the think-aloud process, accompanied by a quick warm-up exercise (not using the dashboard), so as to allow participants to become familiar with the think-aloud procedure. After the warm-up exercise, participants were allowed to browse the dashboard interface for strictly 5 min. After the 5 min had elapsed, the main task began; this consisted of the following four activities:

1. Activity 1. How many times on average was the first video from the first week/module viewed by students?
2. Activity 2. What percentage of people passed the course?
3. Activity 3. How many students finished the peer assessment for <exercise x>?
4. Activity 4. How many people made a forum post on <date x>?

Each activity allowed the participants to get the answer however they liked, and at the end of each activity, they rated on a 6-point scale, the ease with which they were able to answer the question (6 being strongest agreement). Each activity asked how the participant navigated to their answer, with one of the following:

- “I used the Report Categories”
- “I used the Functional Domains”
- “I used the Navigation Sidebar”
- “Other (Please explain)”.

Upon the conclusion of the activities, a SUS questionnaire was administered, which sought to reveal usability insights to the interface. Lastly, a semi-structured post-interview questionnaire was administered to tease out what participants thought would be required to fully utilize the dashboard, including thoughts on training, features found most/least useful, and additional elements they would like added to the dashboard to enrich their and others experience.



### 6.5.2 Apparatus

The machine used for each session was a regular 13-inch MacBook, with a screen resolution of  $1280 \times 800$  pixels. This apparatus was chosen, given the near ubiquity of laptops used around campus, its portability, as well as to situate the experiment within common hardware, for example compared to a dual or triple screen setup which is less common in the university workplace, and which would have unduly added bias to the interface exploration. Each participant was also provided with a computer mouse, which they could use in-lieu of the trackpad, if they so wished.

### 6.5.3 Participants

Eleven participants were sourced from UNSW Australia, who had not been exposed to the MOOC dashboard previously. Each session was run individually by one interviewer. For the think-aloud components, minimal help was provided (e.g. if the participant managed to close the web-browser, this was re-opened for them). The age range was 26–54 with a mean age of thirty-seven. Participants had a diverse background set, drawing on a pool of academics, project managers, as well as educational support staff. Four participants had previous use with dashboards.

Overall, participants rated themselves as above average with regard to the use of computers and average with ease of use with new technology. Participants reported that their visibility to teaching data had mainly been used on an ad hoc basis, either for personal research projects or for helping strengthen cases for promotion. Most participants commented that they would like more timely access to data if at all possible, to help in both their research and teaching efforts.

## 6.6 Results

Three forms of analysis are reported. First is a study of the outputs from the four activities undertaken in each session. This sought to find out features used as related to the pedagogical design of the dashboard, and also the ease of use with finding information for a person first exposed to the dashboard without prior training from analytical/data team members. Next, the usability of the interface is evaluated, as drawn from a SUS questionnaire. Lastly responses from a post-interview questionnaire are analyzed. In addition, free-form comments within each section were used to support interpretation of the analysis and to learn how participants perceived the experience.

### 6.6.1 Activity Analysis

The analysis of activity was studied in terms of how many participants reported the correct (i.e. expected) response; the time spent toward determining their response; and how the participants arrived at their answer. Table 6.4 summaries these aspects.

For activity 1 (how many times on average was *the first video from the first week/module* viewed by *non-signature track students*?), the profile of ease of completing the task was: 2× strongly disagree; 4× disagree; 1× somewhat disagree; 3× somewhat agree; and 1× agree, for an overall score of 2.72/6, see Fig. 6.9. This question took the longest on average to answer by participants, and was also the question, participants least answered accurately (that is, arrived at the expected response). Participants commented on a number of issues, for example:

‘The expanse of information available made it a slower process than I would have liked when trying to find the information. I assume this would be easier the more I used the tool’

‘Screen size is small, which makes it difficult to read titles’; ‘lots of data on one screen’

‘There is no way to understand what is meant by first video’.

It appears a number of participants had conceptual issues regarding the use of. The word ‘first video’, where review of the think-aloud recordings revealed participants asking themselves, ‘first as in any video of the module watched first, first as in the first video on the page’. The particular question required navigation to the video tab, and hovering over the top-most video within the first week/module, which would have revealed the required answer. Half of the participants reported totals rather than averages (for the purpose of the question, these were marked as accurate). The four people who did not arrive at the correct answer, pulled their answer from un-related tabs which they misinterpreted as being videos, or tried to

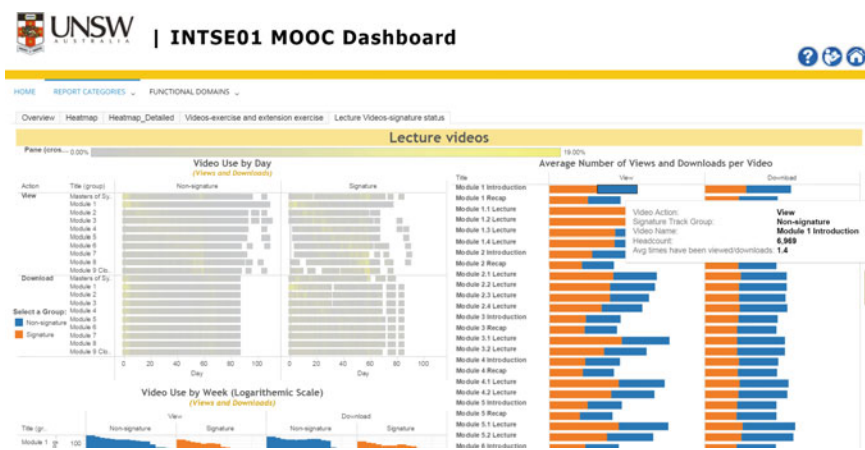


Fig. 6.9 Capture of the panel for Activity 1: the answer is shown in the pop-up box

filter to just the first week based on assumptions of when they thought the first week began. Only 1 participant decided to give up in the end (after 7 min). Of the 11 participants, all but one commented that they would have liked more accompanying documentation; however it was not observed that a single participant referred to the help-menu in the top right corner. This highlights that an on-board process, explaining the features and navigation layout may be needed for future deployments. Figure 6.9 shows a screen capture of the task with the solution.

For activity 2 (What percentage of people *passed* the course?), the profile of ease of completing the task was: 1× somewhat disagree; 5× somewhat agree; 3× agree; 2× strongly agree; for an overall score of 4.54/6. From participant's comments, even though 10 out of 11 people reported the correct response, there were still a small number of participants (4 people) who were confused by the usage of the word 'passed', for example:

definition of passed is not clear

what is the difference between certificate and completed.

Otherwise most people were able to report the response, based on the literal description of passed (as had been named in the dashboard), for example:

the question was clear and it asked for a pretty broad/simple answer it was a fairly easy task.  
obvious to me where to find the answer

the pop-ups were the key to helping me find the information I needed but I ignored them at first.

Half of the comments reported, still asked for further help dialogs or information. A couple of comments mentioned they would have liked 'Report Categories' to be called 'Menu', and 'Functional domains' to be called 'Pie chart', to match their mental models of the dashboard as they perceived it. Figure 6.10 provides the screen capture of the task.

For activity 3 (How many students *finished* the *peer assessment* for <exercise x>?), the profile of ease of completing the task was: 2× strongly disagree; 2× disagree; 2× somewhat agree; 4× agree; 1× strongly agree; for an overall score of 3.63/6. Again participants mentioned that there was:

'Way too much info on screen'/'too much data on one page'.

However, it was mentioned that the pop-up's over graphs were useful, but overall, participants wanted less graphs per page. Figure 6.11 shows the screen capture of the task.

For activity 4 (How many people made a forum post on <date x>?), the profile of ease of completing the task was: 2× strongly disagree; 1× disagree; 2× somewhat disagree; 2× somewhat agree; 3× agree; 1× strongly agree; for an overall score of 3.73/6. For the purposes of this question, the participants needed to direct themselves to the forum section of the dashboard, and click on the forum

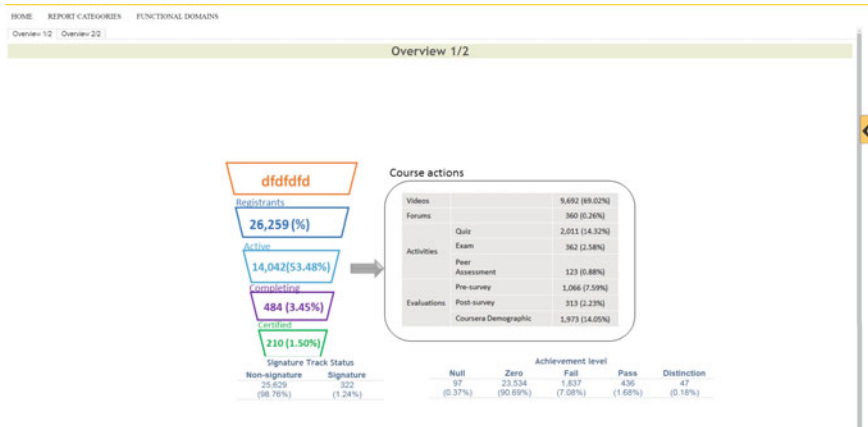


Fig. 6.10 Capture of the panel for Activity 2: the solution is in the table bottom right

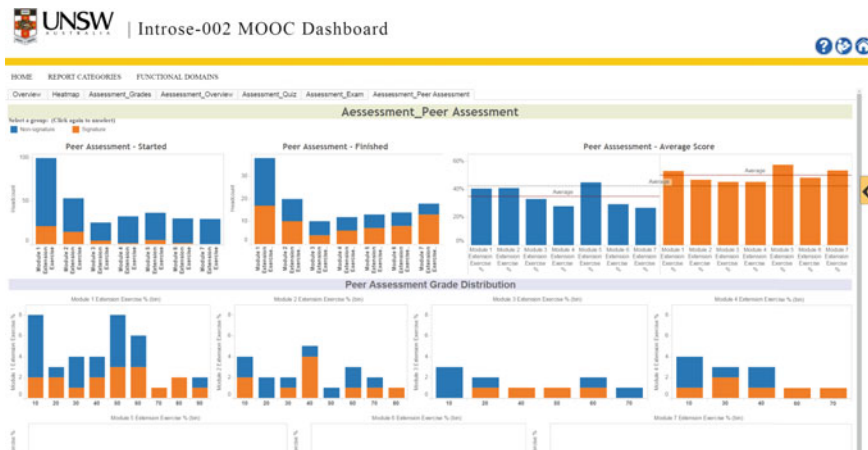


Fig. 6.11 Capture of the panel for Activity 3: answer in the second chart in the first row

graph to reveal the split between posts and comments. Roughly half were able to do this, however, if they reported the combined posts and comments, this was not deducted from the number reported as correct. This resulted in a few participants wondering if this activity was ‘a trick question’. This was again, combined with a similar theme seen in the previous activities of their being a lot of options/ways to navigate to information. Participants also raised concerns about terminology use and jargon they may not be familiar with. Figure 6.12 shows the screen capture.

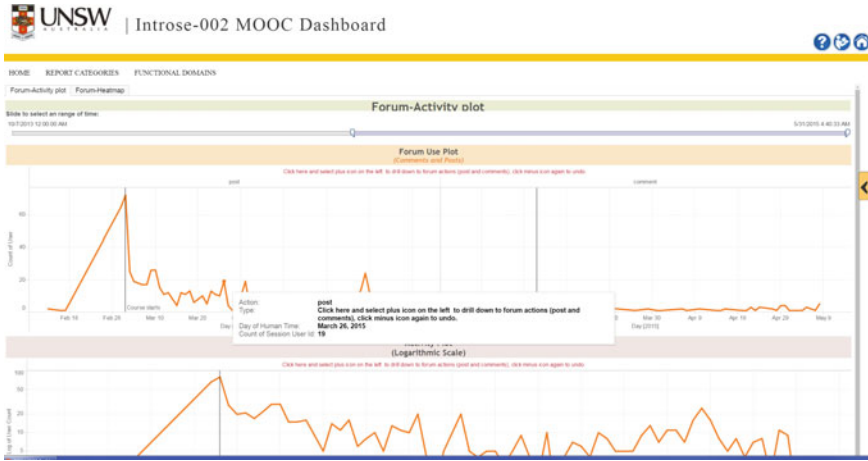


Fig. 6.12 Capture of the panel for Activity 4: answer in the popup box

## 6.7 System Usability Analysis

The SUS usability questionnaire consists of ten questions, with each odd numbered question posed in a positive frame, and each even numbered question posed in a negative way. Thus, for each odd-numbered question, the closer to ‘strongly agree’ the better and for each even-numbered question, the closer to ‘strongly disagree’ is better. The questions in combination can be used to formulate a SUS score, which can help inform an interpretation of the usability of the system under investigation.

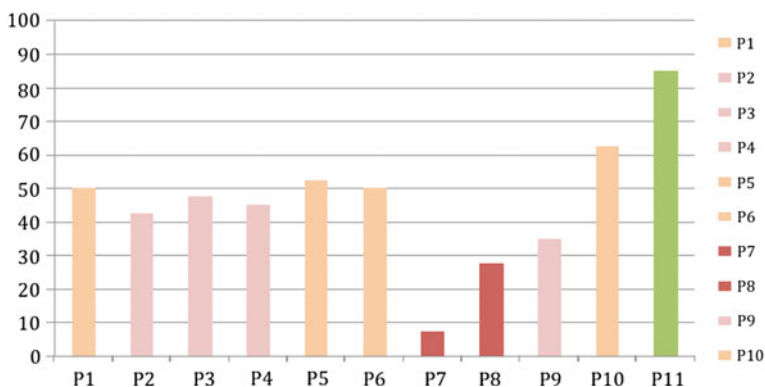
The score itself works on a sliding scale, with 85+ as excellent, 70–85 as good, 50–70 as okay, 35–50 as poor, and under 35 meaning a lot of improvement is required in terms of usability. The profiles of each question are given in Table 6.5. The SUS scores are shown in Fig. 6.13.

Overall, what can be glanced from the participants first use of the dashboard (without any aids, tutorials or prior demonstrations), is that scores largely fell between the ‘poor’ to ‘okay’ range. A number of the free-form comments mentioned wanting to have an initial walkthrough session, and most mentioned they believe their scores would revise upward the more time they spent with the interface. Thus the fact that only two participants registered in the ‘very poor’ category (P7/P8), shows that the initial design, whilst suffering cosmetic issues was largely usable.

The individual questions brought forth that, people would likely continue to use the system, they did not find the dashboard overly technically complex, however they would prefer the dashboard to have less graphs/images per tab/webpage. This was a clear indicator that users found the provided information in its current form as partly overwhelming. Users reported that they likely would not need someone technical to help them out, but did ask for an initial primer with the data team to help with explorations. Overall, most questions, were near neutral in response.

**Table 6.5** SUS scoring table, from <http://usabilitygeek.com/how-to-use-the-system-usability-scale-sus-to-evaluate-the-usability-of-your-website/>

Question	SD	D	N	S	SA	Overall
(1) I think that I would like to use this system frequently	1	1	3	4	2	3.45/5
(2) I found the system unnecessarily complex	0	1	2	6	2	3.81/5
(3) I thought the system was easy to use	1	4	5	0	1	2.63/5
(4) I think that I would need the support of a technical person to be able to use the system	2	5	1	2	1	2.54/5
(5) I found the various functions in this system were well integrated	0	3	6	2	0	2.91/5
(6) I thought there was too much inconsistency in this system	3	4	2	2	0	2.27/5
(7) I would imagine that most people would learn to use this system fairly quickly	1	6	1	2	1	2.63/5
(8) I found the system very cumbersome to use.	0	1	3	6	1	3.63/5
(9) I felt confident using the system	1	6	3	0	1	2.45/5
(10) I needed to learn a lot of things before I could get going with this system	0	2	3	5	1	3.45/5



**Fig. 6.13** Distribution of scores in the SUS

## 6.8 Post-interview Questionnaire Analysis

The semi-structured post-interview questionnaire asked six questions to the participant, the first three were about the think-aloud process, and the later three questions about aspects of the dashboard. Overall, no participants had objections or issues with the think-aloud protocol employed. The last three questions are now reported separately.

### ***6.8.1 What Guidance/Training Do You Think Is Necessary to Use the Dashboard?***

The participants asked that there be more explanations about terms (interestingly many of these explanations were present on the interface, generally within tooltips or question-mark icons, though most participants did not appear to notice these). Some asked for an annotated page describing the functions of the dashboard. Again, this was present in the help icon in the top right, with none of the participants investigating it. Some asked for a video guide to accompany a FAQ section. Participants reported that:

if it was exactly as it is, without any modifications, no training or support documents are necessary just more time to understand the different aspects also information built into the system such as the question marks.

An overwhelming theme that emerged was the want for more hands-on demonstrations of possible ways to use the dashboard, for example:

walked through it at least once to explain what features are where

If you are not very familiar with technology (like a lot of academics are) I think it would be very frustrating to be honest. If you were very comfortable with technology I think you would work your way around it but I think you would definitely need some sort of small group workshops to get comfortable using it, and some sort of support mechanism to go back to

someone from the data team to go through the main elements of the dashboard to explain what data is being displayed and how you would go about getting it. Documentation to walk you through the steps to get it. Would be good to know if the data is something that could be used to develop the next courses.

### ***6.8.2 Which Features Were Most Useful and Why?***

Participants commented on a range of features, for example, the homepage with the report categories and functional domains as it was ‘uncluttered’, and ‘aesthetically the visual representation of students worldwide was good. At least one participant commented on liking the use of the tooltips, as in:

pop up info when hovering over graphs. That was very informative when looking at something unfamiliar that you are not sure about.

Within the same question, a number of participants mentioned the confusion with the use of calendar weeks, compared with say course weeks, for example:

found referring to calendar week number rather than course week number very confusing (i.e. starts at week 31 on forum heatmap page).

Also participants liked that the charts were interactive. For example:

I liked the interactive charts. The plot charts being able to change the dates and that you can hover above things to get info. It just needs to be laid out a bit better because the navigation is a bit clunky and doesn't really do it justice. The more interactive and the more simple it is to get to that stage the better. A pre-set report that you could just click on and it was there would be good. There is a lot of information there, maybe it needs to be split up a bit better and the navigation needs to be sorted but I did like the hovering thing.

### ***6.8.3 Are There Any Additional Features You Would like to See?***

The additional features turned into more of a commentary to a request to reduce the amount of information per tab/webpage. For example:

There is a lot of information on each page, maybe split some of the pages, into separate tabs/pages.

too much info on screen at the moment. Need the option to have just one chart on the screen, i.e. click on it to fill the screen. Or, have the option to build your own by screen selecting which charts you want to compare. Just the individual charts on each screen would be best. Think it is just too messy as it is. If you want to find specific piece of info it is too hard to find I think.

'I would like to see less features in lots of pages.' Only one person mentioned an actual feature request, and that was a 'search mechanism'.

## **6.9 Conclusion and Future Directions**

In this chapter we presented an evolving process aimed to create a sustainable and reusable dashboard for MOOCs, which was intended to provide a tool for a variety of stakeholders to make sense of what is happening in the MOOCs developed and delivered by our institution in two different MOOC platforms. The challenges of the process and the choices made in the implementation have been described in Sect. 6.3.

As the development process relies on an active engagement with stakeholders, the prototyping phase of widgets is a responsive process to target stakeholders' needs. These are then included in the main panels of the dashboard following a principled approach, which relies on a framework informed by both data and reporting needs and affordances. User testing, like in the small scale study presented in the second half of the chapter, drives the process of development demonstrating responsiveness to the stakeholders' needs.

In the description of the implementation process we have established workflows, requiring minimal technical skills to generate a visually pleasing layout with our prototypes. Our framework is a step towards removing difficulties that have



commonly plagued multi-tool adoption and advances the work carried out by others trying to solve similar problems (Seaton et al. 2013, 2014; Verbert et al. 2013b; Qu and Chen 2015; Leon Urrutia et al. 2016).

Based on the feedback from participants, we believe that the framework proposed goes towards the right direction in alleviating the disorientation typical of users beginning to make sense of MOOC user activity and our preliminary experiences are promising and the lessons learnt, in aiding the wider MOOC-related community in their own data related exploration efforts.

As indicated, the process is evolving and it is our intention to continue the developments, clarifying in more detail the elements of the framework by subdividing the functional domains and reporting categories according to different layers of analysis which target specific stakeholders. For example bursts of activity might be caused by problems with learning materials or inappropriate student behaviour: when these become visible to designers or instructors, this can provide an opportunity to counteract and resolve issues quickly. Activity bursts can also be caused by other factors, such as a particular learning design and as such highlight a particular topic or idea of interest for instructors. Academic managers might also be interested in such bursts because they highlight good (or bad) practice, which others should learn from and re-use (or potentially avoid).

There is no doubt that dashboards offer great opportunities for understanding MOOC activity (Qu and Chen 2015; Leon Urrutia et al. 2016) and the effectiveness of the pedagogies implemented in MOOCs; this chapter provides useful and practical observations in order to further the development of data-driven efforts to represent learning-in-action in online learning environments.

In terms of the development, it is our intention to make the dashboard developed for FutureLearn open. On one hand this will allow other institutions to learn from their MOOCs, on the other, it will open up the possibility for further collaborations with others, improving the quality and effectiveness of the visualisations.

## Appendix

Please note that some of the products at the time of publication may not be available as some companies have already been acquired/merged by September 2016.

Coursera	<a href="http://coursera.com">http://coursera.com</a>
Dashzen	<a href="https://www.dashzen.com/">https://www.dashzen.com/</a>
Drillable	<a href="https://www.salient.com/drillable-dashboards/">https://www.salient.com/drillable-dashboards/</a>
Ducksboard	<a href="https://ducksboard.com/">https://ducksboard.com/</a>
FutureLearn	<a href="http://futurelearn.com">http://futurelearn.com</a>
Infocaptor	<a href="http://www.infocaptor.com/">http://www.infocaptor.com/</a>
Klipfolio	<a href="https://www.klipfolio.com/">https://www.klipfolio.com/</a>

(continued)

(continued)

Leftronic	<a href="https://www.leftronic.com/">https://www.leftronic.com/</a>
Logi	<a href="http://www.logianalytics.com/expertise/dashboards-and-reports/">http://www.logianalytics.com/expertise/dashboards-and-reports/</a>
Qlik	<a href="http://www.qlik.com/us/">http://www.qlik.com/us/</a>
Shiny Dashboard	<a href="https://rstudio.github.io/shinydashboard/">https://rstudio.github.io/shinydashboard/</a>
SISENSE	<a href="https://www.sisense.com">https://www.sisense.com</a>
Tableau	<a href="https://tableau.com">https://tableau.com</a>
The Dash	<a href="https://www.domo.com/solution/beyond-business-dashboards">https://www.domo.com/solution/beyond-business-dashboards</a>
Ubiq	<a href="http://ubiq.co/">http://ubiq.co/</a>

## References

- Allio MK (2012) Strategic dashboards: designing and deploying them to improve implementation. *Strategy Leadersh* 40:24–31. doi:10.1108/10878571211257159
- Amo D (2013) MOOCs: experimental approaches for quality in pedagogical and design fundamentals. Proceedings of the first international conference on technological ecosystem for enhancing multiculturalism. ACM, New York, pp 219–223
- Arnold KE, Pistilli MD (2012) Course signals at Purdue: using learning analytics to increase student success. Proceedings of the 2Nd international conference on learning analytics and knowledge. ACM, New York, pp 267–270
- Azevedo R, Harley J, Trevors G, et al (2013) Using trace data to examine the complex roles of cognitive, metacognitive, and emotional self-regulatory processes during learning with multi-agent systems. In: Azevedo R, Aleven V (eds) *International handbook of metacognition and learning technologies*. Springer, New York, pp 427–449
- Baggaley J (2013) MOOC rampant. *Distance Educ* 34:368–378. doi:10.1080/01587919.2013.835768
- Bates T. (2012) What’s right and what’s wrong about Coursera-style MOOCs | Tony Bates. <http://www.tonybates.ca/2012/08/05/whats-right-and-whats-wrong-about-coursera-style-moocs>  
Retrieved from 23 April 2015
- Bayne S, Ross J (2014) *The pedagogy of the Massive Open Online Course (MOOC): the UK view*. Higher Education Academy, New York
- Brooke J (1996) SUS-A quick and dirty usability scale. *Usability Eval Ind* 189:4–7
- Campbell JP, DeBlois PB, Oblinger DG (2007) Academic analytics: a new tool for a new era. *Educ Rev* 42:40
- Churchill D, King M, Fox B (2013) Learning design for science education in the 21st century. *J Inst Educ Res* 45:404–421
- Clow D (2013) MOOCs and the funnel of participation. In: Proceedings of the third international conference on learning analytics and knowledge. ACM, pp 185–189
- Cobos R, Gil S, Lareo A, Vargas FA (2016) Open-DLAs: an open dashboard for learning analytics. Proceedings of the third (2016) ACM conference on learning @ scale. ACM, New York, pp 265–268
- Conrad F, Blair J, Tracy E (1999) Verbal reports are data! A theoretical approach to cognitive interviews. In: Proceedings of the federal committee on statistical methodology research conference. Tuesday B sessions, pp 11–20
- Cooper S, Sahami M (2013) Reflections on Stanford’s MOOCs. *Commun ACM* 56:28–30. doi:10.1145/2408776.2408787

- Cormier D (2009) What is a MOOC. Retrieved from <https://www.youtube.com/watch?v=eW3gMGqcZQc>
- Corrin L, de Barba P (2015) How do students interpret feedback delivered via dashboards? Proceedings of the fifth international conference on learning analytics and knowledge. ACM, New York, pp 430–431
- Corrin L, Kennedy G, de Barba P, et al (2015) Loop: a learning analytics tool to provide teachers with useful data visualisations. Perth, Australia, pp 409–413
- Daniel J (2012) Making sense of MOOCs: musings in a maze of myth, paradox and possibility. *J Interact Media Educ* 2012: Art. 18
- Davis HC, Dickens K, Leon Urrutia M, Sánchez Vera M. del M, White S (2014) MOOCs for universities and learners an analysis of motivating factors. Presented at the 6th International Conference on Computer Supported Education. Retrieved from <http://eprints.soton.ac.uk/363714/>
- Leon-Urrutia M, Cobos R, Dickens K, White S, Davis H (2016) Visualising the MOOC experience: a dynamic MOOC dashboard built through institutional collaboration. In Research Track <https://www.emooocs2016.eu> Graz, Austria, pp 461–471
- Dernoncourt F, Taylor C, O'Reilly U-M, et al (2013) MoocViz: a large scale, open access, collaborative, data analytics platform for MOOCs. In: NIPS workshop on data-driven education, Lake Tahoe, Nevada. Retrieved from <http://groups.csail.mit.edu/EVO-DesignOpt/groupWebSite/uploads/Site/MoocViz.pdf>
- Dominique A (2015) MOOCs, SPOCs, and LAPs: the evolving world of education. In: Huffington Post UK. [http://www.huffingtonpost.co.uk/anton-dominique/moocs-spocs-and-laps-the-\\_b\\_4492046.html](http://www.huffingtonpost.co.uk/anton-dominique/moocs-spocs-and-laps-the-_b_4492046.html). Accessed 12 Apr 2015
- Drachler H, Kalz M (2016) The MOOC and learning analytics innovation cycle (MOLAC): a reflective summary of ongoing research and its challenges. *J Comput Assist Learn* n/a–n/a. doi:10.1111/jcal.12135
- Duval E (2011) Attention Please!: learning analytics for visualization and recommendation. Proceedings of the 1st international conference on learning analytics and knowledge. ACM, New York, pp 9–17
- Fonteyn ME, Kuipers B, Grobe SJ (1993) A description of think aloud method and protocol analysis. *Qual Health Res* 3:430–441. doi:10.1177/104973239300300403
- Fredericks C, Lopez G, Shnyder V et al (2016) Instructor dashboards in EdX. Proceedings of the third (2016) ACM conference on learning @ scale. ACM, New York, pp 335–336
- Govaerts S, Verbert K, Duval E, Pardo A (2012) The student activity meter for awareness and self-reflection. CHI '12 extended abstracts on human factors in computing systems. ACM, New York, pp 869–884
- Kia FS, Pardos ZA, Hatala M Learning dashboard: bringing student background and performance online
- Leon Urrutia M, Cobos R, Dickens K, et al (2016) Visualising the MOOC experience: a dynamic MOOC dashboard built through institutional collaboration
- Mohanty S, Jagadeesh M, Srivatsa H (2013) Big data imperatives: enterprise “Big Data” warehouse, “BI” implementations and analytics Apress
- Nielsen J (1994) Usability engineering. Elsevier
- Pardos ZA, Kao K (2015) moocRP: an open-source analytics platform. Proceedings of the second (2015) ACM conference on learning @ scale. ACM, New York, pp 103–110
- Pauwels K, Amblar T, Clark BH et al (2009) Dashboards as a service: why, what, how, and what research is needed? *J Serv Res*. doi:10.1177/1094670509344213
- Pijera Díaz HJ, Santofimia Ruiz J, Ruipérez-Valiente JA et al (2016) A demonstration of ANALYSE: a learning analytics tool for open edX. Proceedings of the third (2016) ACM conference on learning @ scale. ACM, New York, pp 329–330
- Qu H, Chen Q (2015) Visual analytics for MOOC data. *IEEE Comput Graph Appl* 35:69–75. doi:10.1109/MCG.2015.137
- Rieman J, Franzke M, Redmiles D (1995) Usability evaluation with the cognitive walkthrough. Conference companion on human factors in computing systems. ACM, New York, pp 387–388

- Romero C, Ventura S, García E (2008) Data mining in course management systems: Moodle case study and tutorial. *Comput Educ* 51:368–384. doi:[10.1016/j.compedu.2007.05.016](https://doi.org/10.1016/j.compedu.2007.05.016)
- Ruiz JS, Díaz HJP, Ruipérez-Valiente JA et al (2014) Towards the development of a learning analytics extension in open edX. Proceedings of the second international conference on technological ecosystems for enhancing multiculturalism. ACM, New York, pp 299–306
- Seaton DT, Bergner Y, Chuang I, et al (2013) Towards real-time analytics in MOOCs. In: *IWTA@ LAK*
- Seaton DT, Bergner Y, Chuang I et al (2014) Who does what in a massive open online course? *Commun ACM* 57:58–65. doi:[10.1145/2500876](https://doi.org/10.1145/2500876)
- Siemens G, Gasevic D, Haythornthwaite C, et al (2011) *Open learning analytics: an integrated and modularized platform*. Open University Press
- Stephens-Martinez K, Hearst MA, Fox A (2014) Monitoring MOOCs: which information sources do instructors value? Proceedings of the first ACM conference on learning @ scale conference. ACM, New York, pp 79–88
- Vatrapu R, Reimann P, Bull S, Johnson M (2013) An eye-tracking study of notational, informational, and emotional aspects of learning analytics representations. Proceedings of the third international conference on learning analytics and knowledge. ACM, New York, pp 125–134
- Veeramachaneni K, Halawa S, Deroncourt F, et al (2014a) MOOCdb: developing standards and systems to support MOOC data science
- Veeramachaneni K, O'Reilly U-M, Taylor C (2014b) Towards feature engineering at scale for data from massive open online courses
- Verbert K, Duval E, Klerkx J et al (2013a) Learning analytics dashboard applications. *Am Behav Sci* 57:1500–1509. doi:[10.1177/0002764213479363](https://doi.org/10.1177/0002764213479363)
- Verbert K, Govaerts S, Duval E et al (2013b) Learning dashboards: an overview and future research opportunities. *Pers Ubiquitous Comput* 18:1499–1514. doi:[10.1007/s00779-013-0751-2](https://doi.org/10.1007/s00779-013-0751-2)
- Vigentini L, Zhao C (2016) Evaluating the “student” experience in MOOCs. Proceedings of the third (2016) ACM conference on learning @ scale. ACM, New York, pp 161–164
- Wharton C, Bradford J, Jeffries R, Franzke M (1992) Applying cognitive walkthroughs to more complex user interfaces: experiences, issues, and recommendations. *Proc SIGCHI Conf Hum Factors Comput Syst*, pp 381–388