APPLYING LEARNING ANALYTICS TO CURRICULUM REVIEW: EXAMPLES FROM PROGRAMS

Christine Armatas and Christine Spratt
Hong Kong Polytechnic University
Hong Kong

Abstract

Learning analytics (LA) are used in higher education for predicting student grades or identifying students at risk (Gašević, Dawson, Rogers & Gasevic, 2016). However, there is little research on its use for curriculum evaluation (Méndez, Ochoa, & Chiluiza, 2014). This paper describes a LA-based curriculum review that includes assessment of subject grades, student satisfaction and cohort comparisons. Our results show that using LA as part of curriculum review can provide insights not possible with traditional curriculum review methods and can yield useful and actionable insights. But the challenge remains to develop tools that can assist teachers to conduct LA independently.

Background

This paper reports on how using learning analytics (LA) for curriculum review at the program level can provide insights not possible with traditional curriculum review methods. Most research applying LA in higher education has focused on academic success and retention (Siemens, Dawson & Lynch, 2014), rather than it as an approach to program[1] curriculum review. Historically, curriculum review in higher education has taken a fairly standard approach—stakeholders, usually students and faculty, are surveyed and/or interviewed, standard course and subject performance data are collected, an evaluator (often external to the program under evaluation) is appointed to undertake the review, analyse the data and generate a report with recommendations for improvement. A review of the literature on curriculum review and the current status of the use of LA in higher education shows that LA specifically for curriculum review purposes is under-explored but has considerable potential (Komenda et al., 2015; Méndez et al., 2014; Toetenel & Rienties, 2016). This paper presents applications for LA as part of curriculum review at the program level that clearly demonstrate its usefulness in providing actionable insights that are either not easily obtained or not possible with traditional curriculum review approaches. The progress we have made towards formalizing our approach so that it can be applied to curriculum review of programs more generally and work we have done addressing the challenge of making LA-based curriculum review accessible to teachers through tools that analyse and visualize program or subject data is also briefly discussed.

Curriculum Review and Learning Analytics

Curriculum Review

A curriculum consists of the proposed aims, objectives, learning outcomes and disciplinary content of an educational program. It should be designed considering characteristics of the students entering the program, and have learning outcomes, performance outcomes or competencies that are clear, measurable and reflect the disciplinary requirements of graduates. There also need to be descriptions of the intended pedagogical or teaching and learning approaches (such as active learning strategies, supervision, work-integrated learning, laboratory teaching, e-learning, etc.) and of the substantive curriculum content, along with a robust assessment approach that directs learning and measures intended performance and learning outcomes. There should also be transparent program-wide continuous improvement and evaluation processes.

Internationally, rapid changes in technology and increasing employer demands have come to influence curriculum development and evaluation processes (Cleaver et al., 2017). In higher education, standard approaches to curriculum design generally include those key considerations noted above. Further, curriculum review as a quality improvement strategy has generally included what has been referred to as “the usual incremental and risk-based continuous enhancement processes” (Cleaver et al., 2017, p. 146). Such ‘enhancement processes’ generally seek evidence from stakeholders, principally students and faculty, about their experience of the curriculum and evidence *from* and *about* stakeholders of their learning development against the intended outcomes of the curriculum. At our university, curriculum review approaches include student and faculty surveys of satisfaction with subjects, programs and teaching practices, surveys of the first-year experience and student perceptions of their program after graduation, as well as performance data such as subject and program-level pass rates, progression and program completion data. Quantitative data such as this is normally complemented with qualitative data from documentary analyses and student and faculty interviews.

Creswell and Clark (2017) suggest that the addition of mixed-methods research to review approaches offers strong outcomes-focused evaluation options. Furthermore, the application of LA for curriculum review should not only complement existing methods, but can also add considerable power through its predictive capabilities. While the field is changing, curriculum review in university settings is almost always post-hoc, in that review data are collected at the completion of units of study and programs. Recommendations from the review are then implemented for the next iteration.

Universities collect considerable and varied data about their students across the course of their studies. For example, data relevant to students’ learning behavior are held in student administration records, the learning management system, the library, IT services and other sources. Unfortunately, data often exists in silos and are rarely aggregated, analysed and applied to specific curriculum questions. However, when data are available for analysis as part of curriculum review, the potential exists to provide valuable insights for action.

By using LA for curriculum review, data from multiple sources can be aggregated and analysed and complex and unstructured data can be turned into actionable information about what is happening in the curriculum and how to address performance challenges (Daniel, 2015).

Learning Analytics

A commonly accepted definition of LA is "the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs" (Siemens, 2013, p. 1382). LA can be used to predict student performance, to understand the causes of at-risk learning behaviors and student attrition and for assessing institutional performance (Greller & Drachsler, 2012). Student characteristics, grades, behavior and effort have been found to be highly predictive of past students' success. Furthermore, using this information to help current students with their studies has been shown to improve student grades significantly and decrease fail rates (Mattingly, Rice, & Berge, 2012).

About the Current Work

Our Centre supports academic departments in the implementation of outcome-based education and program review. As part of this work, we have been exploring the application of LA to a range of different data, including institutional survey data and student performance data. Since the graduation of the first cohort of a new four-year curriculum introduced in the 2012-13 academic year, there has been considerable interest in program review to ensure that the curriculum is achieving the intended learning outcomes. There has also been a focus on how to improve programs from a structural (e.g., timetabling) and learning design perspective. Our Centre has worked with program and subject leaders to assist them with curriculum review, adopting an LA approach to complement and extend traditional methods. Through this work, we have developed and tested analytics to address questions raised as part of subject and program review. As discussed next, these analyses can provide insight into program and subject difficulty, relationships between subjects in a program, assessment mix and differences between student groups.

Analytics for Curriculum Review

The approach we have adopted to curriculum review has been to look for analyses that address specific questions that the curriculum review wishes to address. In addition, we have also looked at the data available for analysis as part of the review to determine what insight could be provided relevant to the curriculum review. In the next sections we present a selection of analyses that address important questions about the curriculum that are not easily answered using traditional approaches to curriculum review. At the program level, these questions relate to defining and measuring how difficult a program is as a whole, how student grades are related to student satisfaction, whether the learning outcomes and assessment mix are appropriate and if there are differences on variables of interest, such as grades or satisfaction with the program, between identified student groups. The examples provided are from actual program reviews; however, data is anonymized.

Assessing Program Difficulty

A question of interest in one program review we conducted related to the appropriateness of the level of challenge across all subjects, which we interpreted as the difficulty of the program. While student grades or grade point average could be used to address this question, our approach was to conduct Rasch analysis to compare subject difficulty against students' ability to determine the overall difficulty of the program. A full explanation of Rasch measurement is beyond the scope of this paper. However, one of its advantages is that Rasch analysis can calibrate the person estimates (ability) and the item estimates (difficulty) on the same unidimensional scale (Bond & Fox, 2015). In our case, an "item" represents a subject. For Rasch analysis, the difficulty of a subject was estimated from all of the grades that students who took the subject received, while student ability was determined from their performance in all subjects. Another benefit of conducting Rasch analysis to determine students' ability and subject difficulty is that both use the same scale units, logits (log odds units), which are linear and can be compared on the one scale. When a student's location (ability) on the unidimensional scale is equal to the difficulty of getting a certain grade in the subject, the student has a 50% probability of obtaining that grade. Figure 1 shows the item-person map for Rasch analysis for one program we reviewed. As this figure shows, for this program the mean student ability is above the mean subject difficulty, which indicates that, overall, the program could be more challenging as the average difficulty level of subjects is below the ability of the average student.

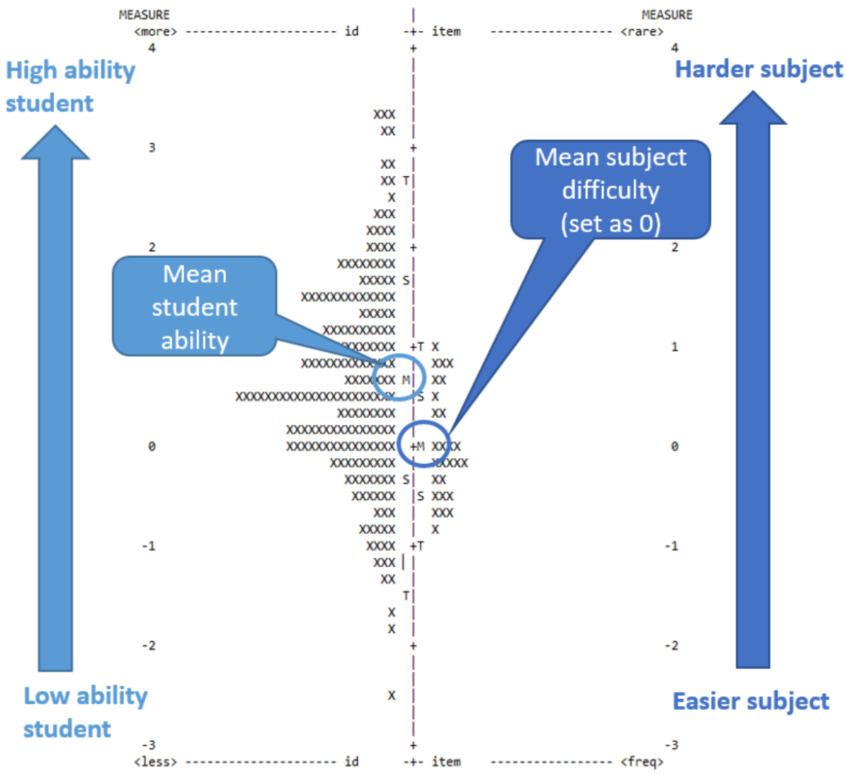

Figure 1. Item-person map showing student ability against subject difficulty. On the right hand side, one x equals one subject (N=27), while on the left hand side; one x equals three students (N=300).

The subject difficulties determined by Rasch can also be used in other statistical tests. An example of once such application is provided next.

Identifying Subjects That Need Revision

Programs have a large number of subjects that students take. Thus, decisions about what subjects need revision can be difficult. Figure 2 shows subject difficulty scores plotted against the subject's mean evaluation score on the end of semester evaluation of teaching survey. In the figure, the lower the subject number, the earlier in the program the subject is taken by students. As shown in the figure, subjects 1, 2 and 4, which are taken by students in the first semester of their first year, receive quite low subject satisfaction scores and are relatively more difficult. Furthermore, subjects 7, 14 and 16 are relatively easier but also have lower satisfaction scores. This analysis and visualization suggest that review of these subjects (i.e., subjects 1, 2, 4, 7, 14 and 16) to determine how to improve satisfaction and/or student performance is needed. In this way, the analysis provides the curriculum review team with information to identify and prioritize subjects for revision as part of the review.

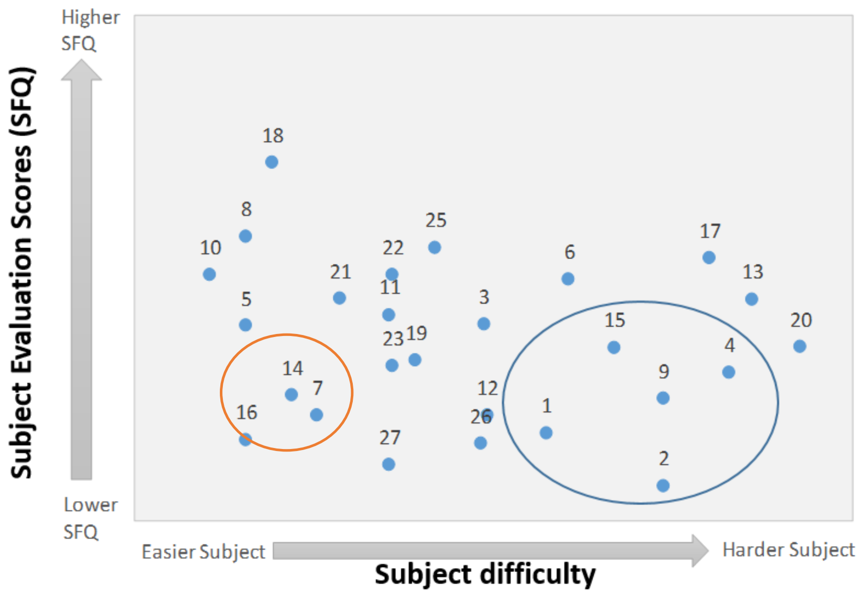

Figure 2. The relationship between subject difficulty and student satisfaction with the subject. Subjects are numbered according to the order in which students complete them in the program.

A second analysis and visualization that provides information about subjects in the program is shown in Figure 3. In this figure the thickness of the line joining nodes in the diagram indicates the strength of the correlation between subjects – the thicker the line the stronger the correlation. From the figure, students' performance in ABC2006 and ABC3001 is strongly correlated with their performance in ABC4001. Similarly, performance in ABC3004 is strongly correlated with both ABC2004 and ABC4006, the latter being low on student satisfaction and low on subject difficulty. Based on this analysis, revision of the lower level subjects (i.e., ABC3004 and ABC2004) to increase their difficulty could be undertaken to better prepare students for the upper level subject ABC4006. What revisions need to be made to achieve this would be determined by further analysis and review of the subjects themselves.

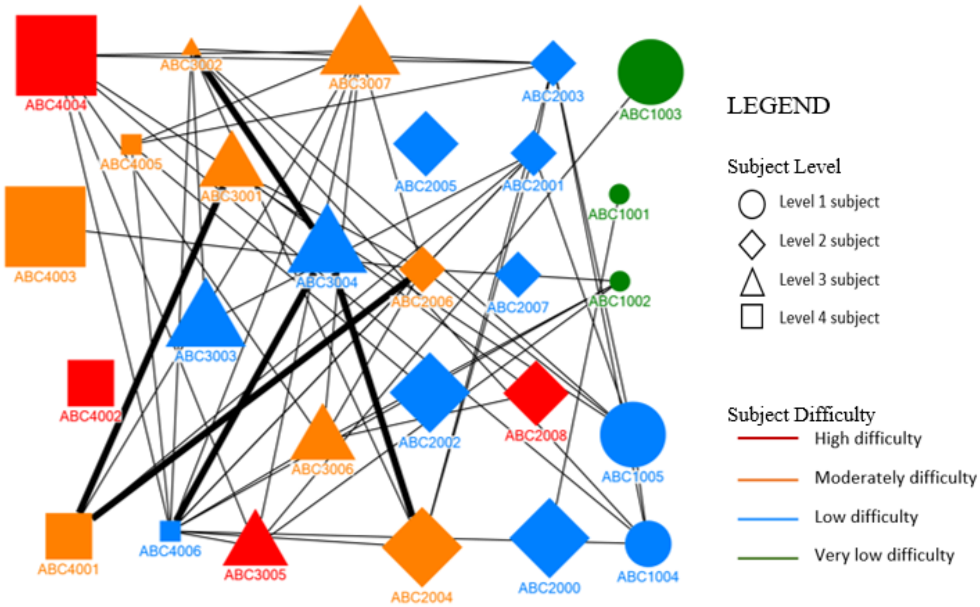

Figure 3. Visualisation of the relationship between subjects together with subject level (indicated by node shape), average student satisfaction with the subject (indicated by node size – the larger the node, the higher the average satisfaction rating for the subject) and subject difficulty (indicated by node colour – see legend for details).

Classifying Learning Outcome Levels and Identifying Assessment Mix

Another consideration in curriculum review relates to learning outcomes and assessment tasks. Figure 4 shows the distribution of learning outcomes for each year level of a program, which have been classified according to Bloom's Taxonomy (Krathwohl, 2002) on the basis of the adjectives in the learning outcome statement. Categories are arranged starting with lower order skills (Remembering) and progressing through to higher order skills (Creating). As Figure 4 shows, for this program, there is a heavy focus on remembering in the first year subjects, but this decreases in both the second and third years of the program. Furthermore, higher level learning outcomes such as Creating and Applying increase from Year 1 to Year 3.

This visualization is useful for checking that the different levels of learning outcomes are distributed appropriately across the program thereby demonstrating a developmental progression. As Figure 5 shows, it is also useful to do this for the different sections of a program (e.g., core, elective, discipline streams, etc.) – for the program shown in Figure 5, that learning outcomes relating to 'Analysing' do not appear in subjects in either of the themes and the core subjects do not have learning outcomes classified as 'Applying' is something that a curriculum review team might wish to address.

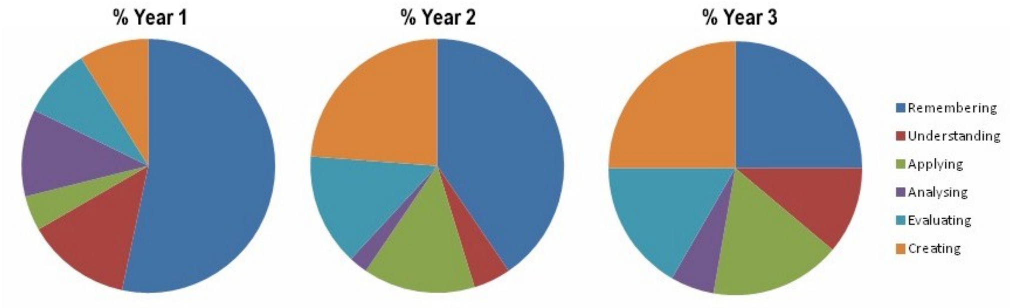

Figure 4. Relative proportions of different levels of learning outcomes classified according to the six levels of Blooms Taxonomy (Krathwohl, 2002) across a program.

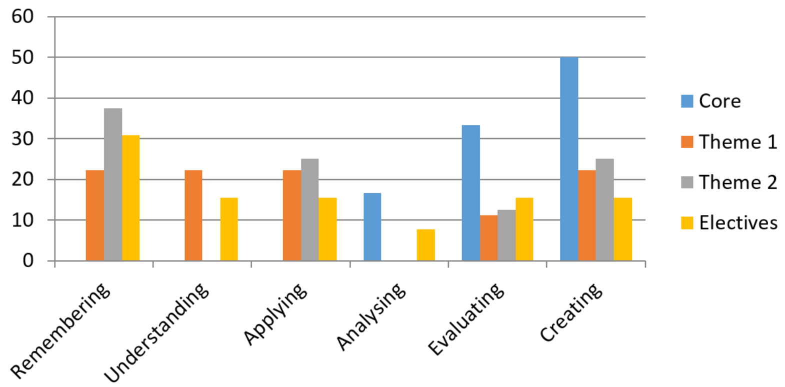

Figure 5. Distribution of learning outcomes by level across different sections of the program.

As well as ensuring that the learning outcomes for the program are appropriate, it is also useful to look at the mix of assessment tasks. Figure 6 exemplifies an assessment mix across a program. Depending on the program or learning approach, some assessment types may need to feature more prominently than others – for example with problem-based learning, a higher proportion of assessment that is project-based might be appropriate. This analysis helps visualize the assessment mix to ensure it is appropriate for the program. Similarly, analysis of the proportion of group versus individual assessment tasks could be used to check the accuracy of feedback from students or staff suggesting there is too much group assessment. A similar analysis and visualization can be produced to check that the relative contribution each assessment type makes towards the student's overall performance on the course is appropriate and that one assessment type (e.g., examinations) does not contribute disproportionately to students' grades in the program.

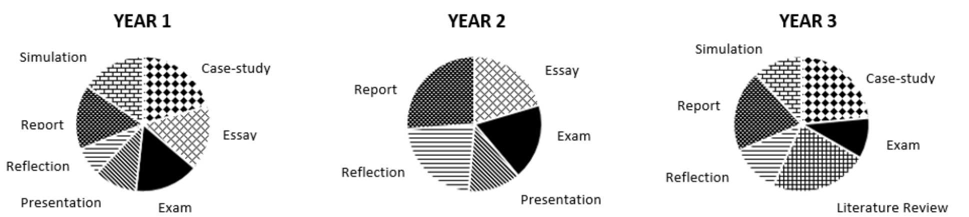

Figure 6. Frequency of assessment types across the program.

Comparing Groups of Students

At our university, programs often have two defined cohorts. The first consists of students who start the program in first year and complete all four years of the program (FY). The second is a group of students who enter the course in third year having received credit based on prior learning. This second group of students, who are referred to as 'senior year admitted students' (SY), often experience challenges because of their entry point to the program and the workload they need to take on to complete the degree in two years.

Our program review approach includes analyses comparing SY students with those admitted to the first year of the program. Figure 7 shows the comparison of these two groups of students in one program on the difficulty of achieving grades for subjects at different levels across the course. As can be seen from Figure 7, the SY students tend to outperform FY students, particularly in those subjects that are typically taken in the first two years of the program. However, by the end of the program, the FY students have caught up and are performing at similar levels. This analysis is useful because it can tell program leaders where students are likely to be struggling so they can provide appropriate support. It also provides reassurance that all students reach a similar level by graduation.

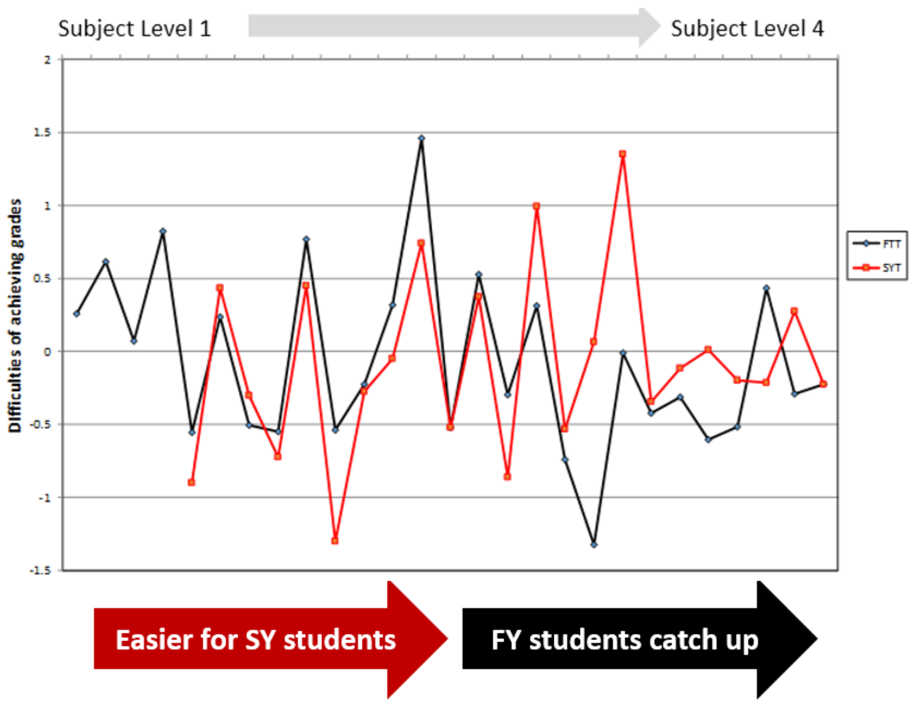

Figure 7. Average subject difficulty across the program for students admitted in the first year of the program compared to those admitted in the third year (senior admitted students). Each data point is the difficulty rating for a subject in the program, with subjects ordered according to when they are typically taken in the program, with Level 1 subjects appearing to the left of the figure and Level 4 subjects to the right.

Discussion

The examples provided above show the usefulness of adopting an LA approach to program review. In each example, the analysis provided insight into the program above that obtained with traditional curriculum review methods. Importantly, the analyses provided findings that were actionable. In our experience, program leaders have found actionable insights most useful; they have commented to us that our analyses provide concrete evidence for their 'hunches' so they can implement strategies for improvement. Our analyses have also told them things about the program that they would not have otherwise known. While feedback from our colleagues has been positive and supportive, we want to formalise our approach. To do this we intend to map data types to analyses that address specific curriculum review questions and detail related strategies for analysis, visualization, interpretation and reporting when conducting LA-based program review.

We have now gained sufficient experience to formalize our LA approach and to test it by conducting further program reviews. However, a key challenge to using LA is having access to relevant institutional and program-level data. Sometimes this is because the data exists in silos and bringing it together for analysis is difficult due to institutional constraints, such as data "ownership." In our situation, data that would be useful for program (and subject) review has not been collected consistently or is not available to our Centre. One of the review recommendations then becomes creating a data collection plan that maps to a curriculum review framework and ensuring that the data is collected for future evaluation exercises.

A second challenge is empowering academics to conduct LA curriculum reviews independently. Many of the analyses we have conducted as part of program review require specialist skills and knowledge that many academics do not have. To address this, we are developing and piloting tools that do the analysis and visualization automatically and include interpretation guides and reporting templates. To date we have a prototype tool that can be used for predicting student performance, visualizing the relationships between variables and comparing groups. We are now documenting our LA approach to program review and refining the tool, with a view to evaluating its usefulness for conducting future program reviews.

Notes

1. By program we mean the timespan of an undergraduate degree. In our context, it is four academic years, that is eight academic semesters where a semester is generally 13 weeks.

References

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch Model: Fundamental measurement in the human sciences*. New York, NY and London, United Kingdom: Routledge.

Cleaver, E., Wills, D., Gormally, S., Grey, D., Johnson, C., & Rippingale, J. (2017). Connecting research and teaching through curricular and pedagogic design: From theory to practice. In B. Carnell & F. Fung (Ed.), *Developing the higher education curriculum (pp.145-159)*. London, United Kingdom: UCL Press. doi:10.14324/1119781787350878

Creswell, J., & Clark, V. (2017) *Designing and conducting mixed methods research.* London, United Kingdom: Sage.

Daniel, B. (2015). Big data and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology, 46*(5), 904-920. doi: 10.1111/bjet.12230.

Gašević, D., Dawson, S. ,Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *Internet and Higher Education, 28.* 68-84. doi: 10.1016j.iheduc.2015.10.002

Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational Technology & Society, 15.* 42-57.

Komenda, M., Vita, M., Vaitsis, C., Schwarz, D., Pokorna, A., Zary, N., & Dusek, L. (2015). Curriculum mapping with academic analytics in medical and healthcare education. *Plos One, 10*(12), e0143748. doi:10.1371/journal.pone.0143748

Krathwohl, D. R (2002): A revision of Bloom's Taxonomy: An overview, *Theory Into Practice*, *41*(4), 212-218.

Mattingly, K. D., Rice, M. C., & Berge, Z. L. (2012). Learning analytics as a tool for closing the assessment loop in higher education. *Knowledge Management & E-Learning: An International Journal, 4*(3), 236-247.

Méndez, G., Ochoa, X., & Chiluiza, K. (2014). Techniques for data-drive curriculum analysis. In *LAK '14 Proceedings of the Fourth International Conference on Learning Analytics and Knowledge. Indianapolis, USA.* (pp. 148-157).

Siemens, G. (2013).Learning analytics: The emergence of a discipline. *American Behavioral Scientist, 57*(10), 1380-1400.

Siemens, G., Dawson, S., & Lynch, G. (2014). *Improving the quality and productivity of the higher education sector – Policy and strategy for system-level deployment of learning analytics*. Canberra, Australia: Office of Learning and Teaching, Australian Government. Retrieved from http://www.olt.gov.au/system/files/resources/SoLAR_Report_2014.pdf

Toetenel, L., & Rienties, B. (2016). Analysing 157 learning designs using learning analytic approaches as a means to evaluate the impact of pedagogical decision making. *British Journal of Educational Technology, 47*(5), 981-992. doi:10.1111/bjet.12423

Author Details

Christine Armatas
christine.armatas@polyu.edu.hk

Christine Spratt
christine.f.spratt@polyu.edu.hk