A Cognitive Processing Framework for Learning Analytics

Andrew Gibson
Science and Engineering
Faculty
Queensland University of
Technology
Brisbane, Australia
andrew.gibson@qut.edu.au

Kirsty Kitto
Science and Engineering
Faculty
Queensland University of
Technology
Brisbane, Australia
kirsty.kitto@qut.edu.au

Jill Willis
Education Faculty
Queensland University of
Technology
Brisbane, Australia
jill.willis@qut.edu.au

ABSTRACT

Incorporating a learner's level of cognitive processing into Learning Analytics presents opportunities for obtaining rich data on the learning process. We propose a framework called COPA that provides a basis for mapping levels of cognitive operation into a learning analytics system. We utilise Bloom's taxonomy, a theoretically respected conceptualisation of cognitive processing, and apply it in a flexible structure that can be implemented incrementally and with varying degree of complexity within an educational organisation. We outline how the framework is applied, and its key benefits and limitations. Finally, we apply COPA to a University undergraduate unit, and demonstrate its utility in identifying key missing elements in the structure of the course.

Categories and Subject Descriptors

K.3 [**Computers and Education**]: General

Keywords

Bloom's revised taxonomy, learning analytics, cognitive processing, curriculum design

1. INTRODUCTION

Learning analytics (LA) is understood in this paper to be "measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs" [15]. This field is undergoing a rapid expansion and a diverse range of approaches have started to emerge [8, 17, 14]. Implementation of LA ranges from basic aggregation of existing data[12], through to course restructuring to work with a specific approach [6]. While using LA simply as a dashboard to report upon student grades and trace contributing factors is relatively easy to implement, it can add little additional value to learning beyond quantitative feedback. In contrast, a LA based restructuring of courses can be highly disruptive to existing teaching and learning, requiring the extensive re-design of units, with the benefits not always proven or immediately apparent to the stakeholders. So both of these approaches risk alienation of the key stakeholders in the learning process, and potentially result in a rejection of the system by those that it was intended to support.

Kruse and Pongsajapan [10] have noted that LA is often focussed on identifying those students that are 'poised to fail', rather than providing benefits for all students. Campbell [4] extends this criticism, claiming that profiling (or placing students in generalised categories) is a common use for analytics, and importantly questions whether this is appropriate.

Despite the best of intentions, latent issues frequently arise in implementations of LA, particularly where metrics are used for judgements about learning. Shum et. al. [13] stated that "a marker of the health of the learning analytics field will be the quality of debate around what the technology renders visible and leaves invisible, and the pedagogical implications of design decisions, whether the design rationale is explicit or implicit." Dietz-Uhler and Hurn [5] question the frequent unintended assumption that metrics extracted from Learning Management Systems (LMS) are a 'proxy for learning'. Beer et. al. [2] identify issues with bringing a managerial mindset to LA: "assumptions embodied by managerialism may be an inappropriate foundation for the application of learning analytics into tertiary learning environments." Assessment and policy theorists warn that the use of learning data has significant implications as a technology of governance and control [11]. Even more concerning are the implications of Campbell's law, which points out that the more that "any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures, and the more likely it will be to distort the social processes it is intended to monitor" [16].

Bach [1] identified the main limitation of deploying learning analytics as "the reliability and validity of the learning outcomes and learner characteristic data used in the models." Hidden problems with validity of data can become more significant, when an LA implementation doesn't take into account the wide and varied needs of those using the system.

For the benefit of learning and greater effectiveness of LA, the design of LA systems needs to address these concerns. In particular, LA needs to work for improvement in learning outcomes through the quality of the information, not just for an increase in the quantity of information.

2. THE COPA FRAMEWORK

We have addressed some of these issues through the development of **COPA**, the **C**ognitive **OP**eration framework for **A**nalytics. COPA is a theoretical framework that is drawn from the cognitive processing dimension of Bloom's Revised Taxonomy[9], a respected educational taxonomy which was developed for the purpose of stating curriculum objectives in terms of levels of learner cognition[3]. This taxonomy provides a widely respected and pedagogically sound foundation for our framework, however our approach could allow for other taxonomies to be used instead.

In developing the framework, our focus was on capturing data that represents the level of cognitive processing required by a learning activity. Levels of cognitive processing can be represented as verbs, or cognitive actions, which range from those articulated in broad curriculum objectives through to their use in more specific learning activities. There are 6 levels in Bloom's Revised Taxonomy which are represented as the verbs: *remember, understand, apply, analyse, evaluate, and create*. It was not our intention for the framework to capture other aspects of the learning process such as the knowledge itself, content complexity, non-cognitive behaviours, and artefacts resulting from learning activities.

Fundamentally, COPA provides a consistent approach for mapping course objectives, learning activities, and assessment items, to the 6 levels represented by these verbs. This provides a flexible structure linking objectives to the cognitive demand expected of the learner. The verbs that are mapped, can then be exposed to the Learning Analytics system via a computational mechanism such as xAPI[7].

COPA was designed to be applied to at least 2 educational contexts, to the curriculum, and to learning activities (including assessment). When applied to a curriculum, COPA can provide insight on the various levels of cognitive demand required by a range of objectives. It can enable an evaluation of the degree of consistency across the curriculum, as well as assist with the identification of potential issues in curriculum structure. In this context, COPA can also provide a mechanism to assess the extent to which learners are engaging cognitively at the levels necessary for the curriculum to be effective.

When applied to learning activities, and the extent to which a learner successfully completes them, the framework can provide an indication of the learner's level of cognitive operation. We have assumed a degree of competence in curriculum implementation, and that therefore, a learner's successful participation in components of a course, is evidence of operation at the mapped cognitive level. In this context, COPA can also indicate the extent to which high levels of cognition are expected over time, and the ability to compare differing tasks based on the level of cognitive demand. It can be used to provide highly personalised information to the learner about the extent to which they are engaging cognitively with certain tasks.

Because Bloom's Taxonomy was originally created to assist with the uniform specification of educational objectives, the accepted application has been to break each of the six levels into subcategories of verbs. See table 1 for an example of how subcategory verbs may be grouped in levels. In the revision of the taxonomy, Krathwohl [9] made explicit the idea that there is some overlap between each of the six levels. It is not expected that verbs only belong to one cat-

egory, as sometimes the sense of how the verb is used may be necessary to ascertain the level for the objective being classified. Similarly for COPA, there is not a rigid requirement that verbs belong to specific categories. However, for automated classification, ambiguous verbs may need to be avoided. It is also important to note with respect to levels of cognition, that the learner doesn't cease to operate at one level when operating at another. Necessarily, there will always be cognitive demand for *remembering* as well as *creating* in a course of learning.

Table 1: Bloom's category verbs with example subcategory verbs

Category verbs	Subcategory verbs
Create	Design, Synthesise
Evaluate	Assess, Critique
Analyse	Compare, Classify
Apply	Calculate, Solve
Understand	Interpret, Summarise
Remember	Recognise, Recall

3. USING COPA

A basic application of COPA to a course of study requires a minimum of 2 steps. Firstly, a mapping process extracts verbs in curriculum documents and associates these verbs with the six levels of the taxonomy. Secondly, the result of the first step is used to filter results of the implementation of the curriculum, which maps results of learning activities back to the six levels. The extent to which each of these levels feature, is recorded into the LA system.

For example, take a course objective that states that students should be able to *critique* a paper, and a corresponding assessment item that requires that students *interpret, compare* and *classify* papers as part of the *critque* writing process. If the results for a particular student show that they did reasonably *interpret, compare* and *classify*, but did not perform particularly well with the *critique*, then based on our example mappings in table 1, this student for this activity might score well in *Analyse* and not so well in *Evaluate*. When taking the aggregate information for all students taking the course, it would be possible to see, in terms of cognitive operation, the extent to which an entire class achieves the set learning objectives. Thus, continuing with the example above, if a number of students exhibited the same pattern of behaviour, and evaluation was the set learning objective, then the course could be seen to be failing to achieve its desired outcomes.

Different courses and activities may use different subcategory verbs in their descriptions, however because these subcategory verbs are mapped to the higher level 6 verbs of the taxonomy (which do not change), many different curricular specifications and activities can be compared at the same cognitive processing level. It is this consistency across diverse elements, that allows COPA to be a powerful contributor to learning analytics.

Another application of COPA allows for the analysis of curriculum design. This application involves mapping both high level strategy documents and low level activity documents in terms of their cognitive demand. The 2 different types of documents can then be analysed in terms of their consistency. For example, if faculty curriculum documents

required all students of the faculty to be *Evaluating* and *Creating* (indicated by subcategory verbs associated with these levels), but the activity documents designed by faculty staff rarely required *Evaluating* and *Creating*, then there would be a mismatch between the cognitive operation expected for individual learning activities compared with the higher level faculty expectations.

This application can be most useful in curriculum design, which is highlighted in the following section, and can also assist with the identification of units of study that are not performing (from a cognitive demand perspective) to the standard expected by the faculty. It can provide a way of using LA to identify whether the action matches the rhetoric with respect to cognitive demand.

Finally, COPA also provides for some degree of prediction. Correlating trends of cognitive processing levels for an individual with learned norms, and associating the resultant data with other analytics such as level of activity, attendance, and engagement, would provide rich possibilities for prediction of future levels of success.

4. EXAMPLE APPLICATION

As a proof of concept, we applied COPA to *Science in Context* (SEB101), an introductory unit in a Bachelor of Science Degree course run at QUT for the first time in 2013. The particular unit that we analyse, was part of a complete redesign of QUT's Bachelor of Science degree to fall into line with the Australian Qualifications Framework (AQF) which is a policy that provides a national unification of all education and training qualifications in Australia. Thus, the analysis that follows is specific to the Australian context. However, we anticipate that it can be applied to other tertiary systems by identifying the corresponding pedagogical instruments and assessment maps.

SEB101 is a unit that lays the foundation for an understanding of the theory and practice of science within its broader social, economic and political contexts. However, students spent a large amount of time performing tasks that were explicitly skills based such as writing and oral presentations. This disconnect between the purpose of the overall unit, and the skills based methodology, created a very real tension between the high level Course Learning Outcomes (CLOs) specified in the AQF, and the more specific learning activities and assessments implemented in class. This made this unit an interesting case study, as we were able to use COPA as a tool to identify discrepancies between the intended purpose of the course and what was actually covered. Any mis-matches can be used in future development of the unit so that is more aligned with its intended CLOs.

For the purpose of this exercise, we utilised both the AQF based CLOs as they are represented in QUTs internal design documentation, and a Unit Guide that was made available to the students in week one of the unit. The Unit Guide provided students with all information relevant to the unit, including assessment tasks and the grading criteria that were to be used in assessing them. Both documents have been made available online for interested readers to refer to at: http://www.users.on.net/~kirsty.kitto/assessment.html

Following the COPA approach, we first identified verbs in the faculty course overview documentation (which is based on the AQF), as well as verbs in the unit guide objectives (not including specific assessment task objectives). We then mapped these verbs to the corresponding levels in the taxonomy, which provided us with information on the cognitive demand required for the course as a whole. An example of how subcategory verbs were mapped to Course Learning Outcomes (CLOs) can be seen in Table 2.

Table 2: Course Learning Outcomes were mapped to subcategory verbs for each level

Level	Verb	CLO
Creating	Synthesising	CLO4
Evaluating	Evaluating, Judging, Deducing, Concluding, Discerning[1], Relating	CL03, CL04, CLO5, CLO6
Analysing	Analysing, Interpreting, Investigating	CLO3
Applying	Applying, Solving, Demonstrating	CLO2, CL07
Understanding	Understanding, Explaining, Communicating, Empathising*	CL01, CLO5
Remembering	Knowing, Articulating, Collecting, Gathering, Recording	CL01, CLO3

Secondly, we used COPA to analyse the objectives for the three assessment tasks and the associated criteria from the task marking rubrics, both specified in the Unit Guide. The assessment tasks included a written critique (30%), a group based oral presentation (30%), and a portfolio of smaller work tasks (40%). The portfolio included marks for completing tasks during workshops, an interview with a scientist, an evaluation of three guest speakers, and an essay about the philosophy of science. The analysis of these activities provided us with information on the levels of cognitive operation that were required by the students to complete the learning activities and assessment tasks. An example of how sub category verbs (for Creating) were mapped with assessment tasks are shown in Table 3.

Table 3: Example of mapping subcategory verbs to assessment tasks for creating level

Level	Verb	Task 1	Task 2	Task 3
Creating	create	1	1	
	synthesise	4	1	1
	integrate	1	1	
	imagine	1	1	
	hypothesise	1		
	compel		1	

For each document analysed, we extracted the sub category verbs and assigned each one to a level in COPA. Most verbs were explicitly stated in the relevant text, but for a few the relevant verb was inferred. Some verbs could be treated at a number of different levels in the taxonomy, but the context in which the verb was used was generally capable of resolving this issue. The analysis was based on a simple frequency count of how many times a particular level was addressed by the document under consideration. While such a frequency count initially sounds too simple to be particularly useful, the discussion that follows will show

that such an assumption is misguided.

We took the frequency count data and expressed it as a percentage of the overall verb count for both course documentation and assessment documentation, which provided us with a SEB101 *signature* for each document set. Table 4 shows the extent to which these signatures correlate.

We would expect that high level objectives of a course would match reasonably well with the assessment items of a course in terms of the cognitive demand placed on students. However, we found that in this instance, that was not the case. This was a disturbing outcome, since it implied that the AQF level CLOs were not clearly followed within the implementation of this unit (as represented in the Unit guide, and in the assessment tasks).

Table 4: A Comparison of course and assessment signatures

Level	Course (percent)	Assessment (percent)
Creating	4.2	20.3
Evaluating	33.3	11.6
Analysing	12.5	16.0
Applying	12.5	18.8
Understanding	16.7	18.8
Remembering	20.8	14.5

The application of COPA in this example, allowed us to identify potential issues in the implementation of SEB101 that can be addressed in the future. If a Learning Analytics system was using COPA for the student results of this unit, we would have also been able to analyse student cognitive processing and compared it to what was expected by the course documentation. This presents opportunity for future research on using COPA for LA.

Although this relatively simple application of COPA to SEB101 proved fruitful from a curriculum design perspective, we anticipate that the application of COPA can inform more than just curriculum design. For example, it would not be difficult to extend the framework across all units that a student is undertaking, and to then to perform an additional mapping that was not performed here: a mapping of the assessment results that the student obtains. If a student consistently performed poorly on tasks revolving around, for example, *analysing*, then this could be shown to them as a graph of expected vs obtained analytical capability. Thus, COPA could identify weaknesses in a student's cognitive development, that the student could then attempt to redress in a conscious manner. In a truly tailored online offering such weaknesses could be incorporated into suggestions as to what tasks or further studies might help the student to round out their cognitive development. Thus, an assessment task that was high on the student's 'weak spot' could be identified or flagged for the student as one that might help them to improve in this area.

On a more pragmatic level, it is important to realise that verb usage in objectives incorporates assumptions about the nature of the tasks that the learner will undertake and the extent to which they accomplish them. Any implementation of COPA should ensure that there is clarity around these assumptions, in order that the application of the framework is consistent, and the resultant analytics are an accurate reflection of reality.

5. BENEFITS AND LIMITATIONS

COPA offers a number of benefits, both at a pedagogical level, and in its design led data capture and reporting. Arguably, the primary benefit of COPA is that it is grounded in solid educational theory. COPA ensures that the LA data is viewed through a lens that has persisted in educational curriculum and assessment design for more than more than 50 years, and that there are generally accepted ways of understanding the resultant data. While there is still room for poor judgement in the implementation of COPA, the educationally based approach with a central focus of learner cognition makes these poor judgements less likely.

A second related benefit of COPA is it's focus on the depth of learning that the student is engaged in. While assessment grades provide LA with a quantity of success and failure over a number of tasks, this data does not usually store the depth of learning associated with those tasks. As such it's possible for students to be highly successful in many tasks that never require a higher level of cognition than applying. COPA, on the other hand provides data on the number of times a student is required to operate at a particular level in the successful execution of set tasks.

Many LA implementations co-occur with curriculum design. A significant advantage of COPA is that when incorporated at the design stage, data capture can be automated. If the mapping of verbs is completed with the design, then the appropriate learning records can be pre-designed and instances generated as the learner participates in or completes activities. For example an analyse verb is associated with a written assessment task in design. The LA system stores this associate between the task and the verb, and when the learner successfully completes the task, a learning record is posted to the learning record store with the learners name, the verb and the task. Given that the objectives containing the verbs will need to be articulated at the design stage anyway, there is nothing more to be done when implementing COPA than to ensure that the system stores the relationship between the task and the verb contained in the objective.

A fourth benefit of COPA is that it doesn't require an all or nothing approach. Partial implementation is possible, or it can be added on to existing curriculum without change to the curriculum itself.

Finally, another of the significant advantages of COPA is that it doesn't need to be explicit in the learning activities for it to work. For example, a problem solving activity may involve searching for academic papers with certain topics being key to the problem, but not explicitly stated in the task, such that higher cognitive levels are required to find the topics that are essential to search for to solve the problem. The search behaviour of each student can be captured, and topics that are explicitly mentioned in the problem are mapped to lower levels, whereas topics that are not explicitly mentioned, but are highly relevant, may be mapped to higher levels. In this situation, the topic search history of the learner gives an indication of how they were thinking about the problem.

There are potential limitations of COPA. We have identified two as follows. Firstly, COPA is reliant on other analytics data to make the analytics about cognitive processing meaningful. For example, it is insufficient to know that a learner has created without knowing what that learner has created and how that relates to other creations. So while we don't apologise for focusing purely on cognitive processing

with COPA, we also recognise the importance of it's interdependence on other aspects of learning.

Secondly, COPA assumes an interest in learner cognition beyond what may be inferred by assessment grades. The value of COPA is related to the desire to obtain richer information about cognitive demand and the extent to which the learner meets that demand. There is little value in implementing COPA if student grades is all that is required to meet the needs of the stakeholders.

6. CONCLUSION

In this paper we have attempted to provide a basis for a learning analytics framework which we have called COPA, that is based on levels of cognitive processing. We have found the cognitive processing dimension of Bloom's revised taxonomy to be effective in allowing us to map verbs from learning documents to levels of cognitive operation. With the recent availability of the xAPI, we anticipate that COPA will provide a useful addition to learning analytics, and to that end we have described the process of applying COPA to a first year science unit as part of a Bachelor of Science degree course. Our early findings in working this implementation have provided valuable feedback to the teaching team for that unit, and indicate great potential in course design. Furthermore, we have identified a possibility for reporting these levels of cognitive processing back to a student undertaking a course of study, which could allow for needs based targeted learning activities and assessment task offerings.

We believe that our new framework, COPA, offers significant benefits for a curriculum design led automated capture of data in a LA system. It provides a way to straddle the all too often seen gap that arises between data that is easy to capture but not particularly beneficial to the student, and data that is highly useful to the student as they attempt to become a better learner, but has so far only been gathered via explicit means (e.g. surveys) that are not specific to the context in which the student is currently embedded. Thus, in paying close attention to respected educational pedagogical principles, we have found a way in which LA could become a powerful tool to enhance the ongoing cognitive development of our students.

7. REFERENCES

[1] C. Bach. Learning analytics: Targeting instruction, curricula and student support. In *8th International Conference on Education and Information Systems, Technologies and Applications: EISTA 2010*. July, 2010.

[2] C. Beer, D. Jones, and D. Clark. Analytics and complexity: Learning and leading for the future. In *29th Annual Ascilite Conference*. 1–10, Nov. 2012.

[3] B. S. Bloom. *Taxonomy of Educational Objectives*. Handbook I: The Cognitive Domain. David McKay Company Inc, New York, Jan. 1956.

[4] J. P. Campbell, P. B. DeBlois, Oblinger, and D. G. Academic Analytics: A New Tool for a New Era. *EDUCAUSE REVIEW*, vol. 42, no. 4, 1–10, July/August 2007.

[5] B. Dietz-Uhler and J. E. Hurn. Using Learning Analytics to Predict (and Improve) Student Success: A Faculty Perspective. *Journal of Interactive Online Learning*, 12(Spring), 9–10, Jan. 2013.

[6] A. L. Dyckhoff, D. Zielke, M. Bültmann, and M. A. Chatti. Design and Implementation of a Learning Analytics Toolkit for Teachers. *Educational Technology & Society*, 15 (3), 58–76. 2012.

[7] Experience API Working Group. Experience API. The Advanced Distributed Learning (ADL) Initiative, Apr. 2013.

[8] R. Ferguson and S. B. Shum. Social learning analytics: five approaches. In *LAK '13: Proceedings of the 3rd International Conference on Learning Analytics and Knoweldge*, May, 2012.

[9] D. R. Krathwohl. A revision of Bloom's taxonomy: An overview. *Theory into practice*, 2002.

[10] A. Kruse and R. Pongsajapan. Student-centered learning analytics. *CNDLS Thought Papers*, 2012.

[11] B. Lingard. Policy as numbers: ac/counting for educational research. *The Australian Educational Researcher*, 38(4):355–382, Nov. 2011.

[12] A. Pardo and C. D. Kloos. Stepping out of the box: towards analytics outside the learning management system. In *LAK '11: Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, 163–167, Feb, 2011.

[13] S. B. Shum and R. D. Crick. Learning dispositions and transferable competencies: pedagogy, modelling and learning analytics. In *LAK '12: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 92–101, Apr, 2012.

[14] S. B. Shum and R. Ferguson. Social Learning Analytics. *Educational Technology Society*, 15(3):3–26, 2011.

[15] P. L. Siemens and George. Penetrating the Fog: Analytics in Learning and Education. In *EDUCAUSE REVIEW*, vol. 46, no. 5 pages 1–6, Sept. 2011.

[16] J. Willis, L. Adie, and V. Klenowski. Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform. *The Australian Educational Researcher*, 40(2):241–256, Mar. 2013.

[17] A. F. Wise, Y. Zhao, and S. N. Hausknecht. Learning analytics for online discussions: a pedagogical model for intervention with embedded and extracted analytics. *LAK '13: Proceedings of the 3rd International Conference on Learning Analytics and Knowledge*, April 08-12:48–56 2013.