Improving Teaching and Learning through Low-Cost and Open-Source Learning Analytics

Nancy E Krusen, Associate Professor of Occupational Therapy, Pacific University

Abstract
Analytics is a continuous iterative process of exploration to gather data, discover patterns, and disseminate information. Learning analytics (LA) are evolving methods to collect, analyze, and report data to provide actionable insight into blended and distance education. The primary aim of the paper was to demonstrate the use of low-cost and open source tools to improve teaching and learning with instructors for whom LA is novel. The study merged Scholarship of Teaching and Learning (SoTL) constructivist perspective with simple educational technology to scrutinize how learning works. The naturalistic study included exemplar questions and LA methods for a single course as a case study. Further research is recommended to replicate outcome evaluation and quality indicators for best practice in distance education using LA.

Keywords: Teaching; Learning; Education; Education, distance; Learning Analytics

Introduction
Analytics is a continuous iterative process of exploration to gather data, discover patterns, and distribute information. Business and healthcare analytics investigate data extracted from reports of past performance to gain insight and direct strategic planning. Researchers use healthcare data, information and technology to decide best practice for individuals and organizations. Learning analytics are evolving methods applying the process to education. Current academic trends include blended and online learning courses, herein referred to collectively as distance education. Blended and online components are often incorporated into health care professions education. Instructors routinely analyze and make changes as part of course evaluation, yet instructors are just beginning to evaluate, analyze, and share digital pedagogy in technology-enhanced learning. Long and Siemens (2011) suggested a shift is necessary from gut instinct to evidence-based education, as has occurred in health care practice. They suggested a need to move beyond data evaluation to provide feedback for course and curricular change to facilitate teaching and learning. The intent of the paper was to apply learning analytics (LA) to identify opportunities for change in distance education. The primary aim of the paper was to demonstrate the use of low-cost and open source tools to improve teaching and learning with instructors for whom LA is novel. The secondary aims were to demonstrate how to gather data about what is happening in a course, how to discover patterns in the data, and how to use the information for course modification. The study merged a Scholarship of Teaching and Learning (SoTL) constructivist perspective, using technology to scrutinize how learning works. The author used the terms learner and student interchangeably. The data reported and lessons learned supported program evaluation with subsequent curricular changes through a case study example of a single course.
Improving Teaching and Learning through Low-Cost and Open-Source Learning Analytics

Learning analytics are evolving rapidly as a unique multi-disciplinary field of research. Academic analytics applies business strategies to address goals, capacity building, retention, and graduation, changing organizational culture, and technology (Elias, 2011; Norris, Baer, Leonard, Pugliese, & Lefrere, 2008). Campbell, DeBlois, and Oblinger (2007) described academic analytics as “actionable intelligence,” or the use of information to improve institutional capacity, teaching, and learning to make proactive decisions for processes and outcomes (p.42).

Norris et al (2008) suggested analysis in education shift away from simply reporting numbers toward a process informing evidence-based decision-making in education. LA may apply descriptive statistics, social analysis, and visualization tools for educational practice (Chatti, Dyckhoff, Schroeder, & Thüs, 2012). Social learning analytics applied visual analysis and representation, revealing the process of community group development and social networking occurring in a “participatory online culture” of distance education (Buckingham, Shum, & Ferguson, 2012 p.5).

The focus of LA is on the process of acquiring knowledge. The most 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,” as noted on the website of the First International Conference on Learning Analytics and Knowledge (2011, np). Elias (2011) further delineated LA as evidence-based decision-making, “an emerging field in which sophisticated analytic tools are used to improve learning and education” (np). LA offers more than a recording device for technology-enhanced learning; they are “a means to an end, not the end in and of themselves” (Wagner and Ice, 2012, p38). LA tools collate data, enabling investigator analysis for insight into student behavior and learning outcomes (Lockyer, Heathcote, & Dawson, 2013; Wagner & Ice, 2012). Behavior-based information describing how learners are currently interacting within the learning environment also supports evaluation of context, the distance course design.

Application of Learning Analytics

Student participation in distance education generates extensive data available for examination, akin to a fire hose dispensing water. Anderson (2008) referred to this massive amount of information as data deluge, requiring an adjustment in scientific method. He suggested a process to first examine the data mathematically, followed by a framework to describe context (np). Wickens (2006) advocated applying technology to pursue the Scholarship of Teaching and Learning addressing Technology Enhanced Learning, or SoTEL. SoTEL lays the groundwork for digital pedagogy for which LA is a tool. Chatti et al (2012) proposed a model to categorize areas of research needs and solutions. They suggested four basic research questions in the application of LA:

1. What? What kind of data does the system gather, manage, and use for the analysis?
2. Who? Who is targeted by the analysis?
3. Why? Why does the system analyze the collected data?

Dietrichson (2013) discussed the application of LA to scrutinize student progress within an online course setting. Papamitsiou and Economides (2014) conducted a systematic literature review of empirical evidence regarding LA and educational data mining in practice. The results of their work
indicated four main areas of interest: pedagogy oriented issues, contextualization of learning, networked learning, and education resources. Ma, Han, Yang, and Cheng (2015) used LA to analyze the impact of an instructor on student engagement across sixteen variables, finding a positive relationship between instructor and student behavior.

Learning Management Systems (LMS) such as Moodle and Blackboard, contain the basic tools to track learner participation and progress. LMS can track specific login information for a course, page or assigned task, including date, time of login, the number of times logged in, and the amount of time spent. Logged data show the number of discussion posts and to whom the learner responded. Logged data also include resource access, test scores, quiz scores, survey, and questionnaire responses.

For useful analysis, the investigator must define the target of LA. Researchers may apply LA to individuals, cohorts or programmatic inquiry. Researchers may use LA for course goals, learning objectives, program evaluation, and/or accreditation standards (Olmos & Corrin, 2012a; Olmos & Corrin, 2012b). Examining correlation over time, high level learning and programmatic impact requires an analysis surpassing simply counting the number of discussion posts or examining graphic representations of interaction (Anderson et al, 2001; Wilson, 2015). Olmos and Corrin (2012a, 2012b) described the use of datasets for a content management system in a medical school, identifying specific clinical problems, learning activity outcomes, resources, and case-based learning activities collected in the metadata. Olmos and Corrin (2012b) examined visualization as a process within LA to represent longitudinal clinical education placements. Olmos and Corrin (2012a) also employed LA to outline curricular content. They used results to develop the curriculum and confirm coverage for accreditation. Both publications provide useful insights into LA applications at a programmatic level, with recommendations for future application.

Investigators use a variety of means to visualize or represent the results of data analysis. Many LMS enable spreadsheets exportable into Excel (Dimopoulos, Petropoulou, Boloudakis, & Retalis, 2013). Free and open-access tools and plug-ins are available which may be applied to different LMS. Among them are Learning Analytics Enriched Rubric (Dimopoulos et al 2013), Social Networks Adapting Pedagogical Practice (SNAPP) (Dawson, Bakharia, & Heathcote, 2010), and exploratory Learning Analytics Toolkit (Dyckhoff, 2014; Dyckhoff, Zielke, Chatti, & Schroeder, 2011). There are many open-source and commercial tools to analyze course data. Many programs are time-consuming for educators to learn easily, require technical expertise to manipulate, and are too complicated to apply to teaching and learning questions. The following examples reveal familiar course questions with simple methods for exploration and representative tools for illumination.

Method
As an exemplar, data were drawn from a single graduate level course in occupational therapy leadership at a university in the western United States. The School of Occupational Therapy recently implemented exclusively online courses in the third curricular year. The study focused on individual learners (n=19), the cohort, and the program. All data were archival and de-identified: the study was exempt from review. The author collected and imported data from Moodle. The investigator employed the following tools: Excel for descriptive statistics of learner behavior; SNAPP for discussion forums; and Voyant text analysis for qualitative examination. Excel is a Microsoft Corporation spreadsheet software program. The SNAPP tool creates social networks with data drawn from discussion forums in LMS, including Moodle and Blackboard (Dawson, Bakharia, & Heathcote,
SNAPP is an open-source application, arranging a web-like diagram to support analysis of interactions between participants. Though a user-friendly tool, SNAPP has not kept pace with LMS upgrades. Voyant is a collection of free tools to analyze digitized texts developed by Sinclair and Rockwell (2012) enabling searches for patterns and outliers. Results are intended as a case study illustration to apply free and open-access tools.

Results
The results of the exemplar case study demonstrate ways in which an educator may apply LA to gather data, discover patterns, and use the information to implement change for a learner, cohort or program. Distance education requires learners to maneuver in digital environments to interact with virtual learning objects, resources, and other participants. It is most useful for educators to frame specific questions about teaching and learning. Results represent common what, who, why and how educational questions for course design, individual and cohort learning, and program evaluation answered using low-cost and open source LA tools.

Reports and graphs
Reports and graphs illustrate the utility of LA to examine the use of course resources in this example. The investigator posed the overarching question: What resources are most frequently used throughout the course?

- **What?** The number of times learners accessed each resource, colloquially referred to as the number of hits.
- **Who?** Individual learners or a cohort.
- **Why?** Distance courses contain multiple resources required for student learning: articles, videos, websites, documents, etc.
- **How?** LMS record frequency, length of time spent, etc. for each learner and each resource.

In this example, the cohort was the unit of focus. The author downloaded data from the LMS to an Excel spreadsheet to generate reports of activity (Figure 1). In order of frequency, resources used include documents (two sets totaled), samples of completed assignments, live video discussion, syllabus, and recorded mini-lectures (three totaled). Figure 1 shows low comparative use of mini-lectures. These data may prompt the instructor to ask additional questions about the value learners place on watching recorded video to make informed decisions for keeping, increasing, decreasing or eliminating the resource: Who are the users of the recorded video? Why do learners use recorded video? Did viewers of the recorded videos attend live session? Did viewers return to the recorded video multiple times? How does use of recorded video compare to use of other resources?
In this example, the investigator also asked: *How often do learners access course resources on a daily basis within the first month of the course, including static and interactive resources?* An Excel chart illustrates daily activity by resource during the first month of a course (Figure 2). The tallest peaks represent the number of discussion posts appended to instructor-generated course topics. The mid-height peaks represent an open discussion forum for introductions, general topics, and informal conversations. Peaks and valleys are clear, indicating high activity for required initial posts and response posts with a lull between. The peaks of the formal and informal forum have similar patterns as participants get acquainted with the course and context. It is possible to point to Saturdays by examining the activity valleys. Though still cyclic, the extreme highs and lows of discussion posts are less dramatic after three weeks. The data may suggest a shift from teacher prompt-student response pattern to a more collegial conversation. The data may prompt additional questions social networking or critical thinking.
Because the high activity in formal and informal discussion forums flattens the depiction of the other resources (in Figure 2), it is difficult to see the use of other course resources clearly. The author modified the question: *What resources do learners access within the first month of the course, excluding all discussion forums?* Figure 3 illustrates the modification – daily activity in an Excel area chart, with informal and required discussion forums removed. It is now possible to see high access to resource documents (back), syllabus (middle, dark), and sample assignments (fore). As anticipated during the first two weeks of a course, students often access the syllabus and sample assignments to learn expectations for the course and context. Notable in Figure 3, no students accessed live videoconference with the instructor or recorded videos during the first two weeks of the course, despite a synchronous tutorial at course start-up. Tracking this lack of activity enabled the instructor to remediate. In SoTL terms, the instructor responded to tacit critique of teaching practice. The instructor posted more detailed instructions for live videoconferencing at the two-week mark, with screen shots demonstrating step-by-step instructions. The instructor also created the first recorded video with a posted link. The graph shows rapid and active response to the video resources at the two-week mark. There is a corresponding peak in informal discussion for the same dates on Figure 2, indicating lively conversation.
Figure 3 Activity level by resource excluding informal (Water Cooler) and required (Topic Discussion) discussion forums

Because of the slow adoption of live videoconferencing, the instructor asked: *How quickly do students access a new resource once posted?* An Excel line chart illustrates the access rate after the posting of a course resource, in this case, the recorded video (Figure 4). Numbers are cumulative. There is an apparent rush to view the video within the first twelve hours followed by relative inactivity. Students return to the video after forty-eight hours. Examination of the raw data indicated that students return days and weeks later to view the recorded videos. From these data, the instructor could examine the access behavior further: *What is the correlation between the number of views and student characteristics? Why do learners return to the resource? What is the connection between the number of views and a particular assignment?* The instructor may choose to alter the design of the course to include tutorial instructions for the next cohort. The instructor may also follow the behavior of the cohort across other courses to determine behavior change.
Social network analysis

Social Network Analysis Programs (SNAP) illustrate the utility of LA to examine the interaction of the individual or cohort. The investigator posed the overarching question: *Who interacts with whom?*

- **What?** SNAP records the number of digital interactions.
- **Who?** Individual learners or cohort or program.
- **Why?** Distance courses contain collaborative spaces in which learners engage with each other, often a required part of coursework. SNAP illuminates student participation in a snapshot or as change over time.
- **How?** SNAP transfers discussion forum posts into visual matrices of frequency, connectivity between individuals, directionality of contact, etc.

In this example, the author applied LA to individuals. Figures 2 and 3 (above) described learner interaction in required and informal discussion. Figure 5 is a SNAP of a single required discussion forum, taking place across one week early in a course. The instructor is the facilitator in the example. The size of the dot indicated the relative frequency of contact initiated by a participant. The arrow direction indicated participant directionality; bidirectional arrows indicate reciprocal dialog. This is a typical facilitator-centric social network pattern in which the instructor initiated (and received) most contact.

Knowledge of the type of discussion forum enabled the instructor to draw informed conclusions about interaction and learning. In other circumstances in which there is no moderator, patterns show different learner-to-learner interaction. The researcher must be aware of the context to grasp the information depicted from the data. Examining the pattern, the facilitator is able to find learners whose

![Figure 4 Access rate following posted resource](image-url)
only contact is toward the instructor to encourage broader interaction, or learners who are ready for the facilitator role. The instructor may ask: *Why are some learners so quiet? How does interaction change during the week?* Patterns may show learners monopolizing or disengaging from interaction, a non-responsive instructor, or suggest tedious discussion prompts. Figure 5 clearly shows an isolated learner, indicated with a red circle. From these data, the instructor may want to ask: *Why is the student disconnected?* The lack of participation may be a single incident, a pattern of behavior, or indicative of a greater crisis. The author has experience in which distance students have suddenly ceased posting in a course. Circumstances included illness, pregnancy, and natural disaster. This is an opportunity for the instructor to understand the learner and the context to appreciate the meaning of isolated data points.

Figure 5 Social network analysis. Number of discussion posts
Textual analysis

Textual analysis, alternately entitled content or discourse analysis, illustrates the utility of LA to examine student-generated texts for content mastery and critical thinking. The investigator posed the overarching question: *Do students connect the core concept of ethics with their experiences during clinical education?* The case study course took place after a three-month, full-time clinical practice experience.

- **What?** The process analyzes digital learner-produced documents, including assignments and discussions.
- **Who?** Individual learners or a cohort or a program.
- **Why?** Learners produce writing potentially reflective of changes in thinking.
- **How?** Digital analysis of downloaded student writing generated core concepts and themes through key words and key words in context.

In this example, the investigator applied LA to the cohort and program. Distance courses support opportunities to explore through examination of student-generated texts. Text analysis illustrated word trends for a single forum in which the prompt asked students to discuss an experience of “a conflict between professional values and justice in practice.” Figure 6 indicates the relative frequency of each term by week throughout the ten-week discussion forum, emergent coding generated through Voyant. The data show high-frequency words “ethical,” “think,” “ethics,” “time,” “therapy,” “patient,” and “situation” (Figure 6). The instructor intentionally excluded the high-frequency words from the prompt to avoid leading learner responses. The pattern in the data suggested a likelihood of critical thinking in which learners connected core concepts of ethics (taught earlier in the curriculum) to applied ethics during their clinical practice placement.

Examining Key Words in Context (KWIC) allows a deeper examination of high-frequency word use and their relationship to each other. The instructor asked: *How do students connect the core concept of ethics with experiences during clinical education?* In other means of textual analysis in Voyant, Figure 7 illustrates an example of KWIC. This KWIC reflects the same discussion prompt, without the inclusion of key words or instruction to connect specific supporting documents. The expanded discussion entry in Figure 7 used terms related to ethics, and has specific references to professional documents, “there was something ethically wrong with how his case was being managed and handled. Based on the occupational therapy code of ethics I felt I was experiencing conflict with the principle of beneficence…” This exemplar post is 1 of 329 instances of KWIC containing the high-frequency words, suggestive that students engaged in critical thinking, connecting professional values and justice in practice.
1) I’m glad for sharing this experience and ethical dilemma. It is true, there asked for him to be placed on medical hold. Many things ran through my mind, which initiated several lengthy conversations with my CI. I had many reservations and thought that his placement at inpatient rehab was completely inappropriate and his pain was being inappropriately managed. But above all I had ethical reservations and felt that there was something ethically wrong with how his case was being managed and handled. Based on the occupational therapy code of ethics I felt that I was experiencing conflict with the principle of beneficence, which is “to do good”, and nonmaleficence “to do no harm” (Ledlow

1) I’m glad especially when the team has ethical concerns. Good example! I am

1) I’m glad never really dealt with the ethical dilemma of ceasing treatment because

1) I’m glad on a shared commitment to ethical standards. I am so proud

1) I’m glad weeks ago describing leadership. Ethical was a common thread most

1) I’m glad and Ethics Standards (2010) is ethical action is a manifestation of

1) I’m glad Ledlow & Coppola suggest that. The ethical tone and moral expectations in

Figure 7 Key words in context
Discussion
The intent of the paper was to apply learning analytics (LA) to identify opportunities for change in distance education across learner, cohort, and program. The results of the case study are not intended as generalizable. The study revealed answers to exemplars of four types of research questions about student participation through the application of LA: what? (data and environments), who? (stakeholders), why? (objectives), and how? (methods) (Chatti et al, 2011). The author did not intend the questions nor the LA presented in this article as comprehensive; they serve as examples of possible free and open-source tool applied to readily available data. The study embraced a Scholarship of Technology Enhanced Learning constructivist perspective to scrutinize how learning works in a single case study course.

The secondary aims were to demonstrate how to gather data about what is happening in a course, how to discover patterns in the data, and how to use the information for course modification. LA can reveal student participation through a variety of data and analysis. Reports, graphs, network analysis, and textual analysis illuminate learners’ interaction with course resources, course content, and other participants. LA can target learners and instructors. Social network analysis is particularly useful to examine interaction between these participants. LA can answer specific questions about individual behavior and behavior of the cohort. Ipsative assessment may scrutinize student or cohort evolution over time. Analytics may apply quantitative and qualitative analysis to collected data. Manipulating the data into different visual representations prompts different perspectives, recognition of novel patterns, or formulation of new questions. Few (2009) stated viewing only raw numbers or statistics can “shear away much of the richness and subtlety in data” (p.6). Ellis (2013) suggested LA be included in formative and summative assessment to improve the scope of educational practice, the primary aim of the SoTL. It is necessary to find the meaning of the relationships through the application of LA and metacognitive analysis.

Theories, models, and frames of reference undergird professional education. Each professional program typically describes a unique mission, vision, philosophy, and curricular design. The primary aim of the paper was to demonstrate the use of low-cost and open-source tools to improve teaching and learning with instructors for whom LA is novel. The field of LA research provides food for thought, encouraging day-to-day educators to pursue SoTL through LA. Lockyer et al (2013) advocated the need to align LA with learning, curricular, and course design. LA application could be associated with elements of theories such as behaviorist, cognitivist, humanist, situational or transformational learning. LA application could be associated with profession-specific models. LA may support pedagogy through programmatic evaluation or review of accreditation standards.

Implications
Exploring the literature and conducting the case study suggested opportunities for exploration across several issues during the study.

- **Legal and ethical standards.** Universities have “a moral responsibility for employing information which aims to provide more effective and relevant support” (Slade & Prinsloo, 2014, p299). The authors proposed specific principles as moral and ethical guidelines to address ethics and LA within higher education (Prinsloo & Slade, 2014; Slade & Prinsloo, 2013). The current
legal system is actively refining policies, procedures and practices involving use of student data (Campbell, DeBlois, Oblinger, 2007; Siemans, 2013).

• **Develop pedagogy for best practice.** Seimans (2013) makes a case for LA as a discipline, an academic and research field in its own right, a contentious view. SoTL (and SoTEL) critically assesses educational practice to inform change, investigating theory, practice, pedagogy, and learning. Koh and Chong (2014) and Van Barneveld, Arnold, and Campbell (2012) described LA as operationalizing those concepts. Van Barneveld et al (2012) further described the relationship between LA and SoTL as complementary and symbiotic. The issue is contentious, prompting exploration of existing and new theories, models, evaluation criteria, quality indicators, curricular design, course design, and pedagogy.

• **Involve stakeholders in development.** Engaging students, instructional design staff, and technical developers with faculty in stakeholder-driven inquiry can offer diverse perspectives for diverse needs. Researchers recommend interdisciplinary collaboration to better understand technological and pedagogical issues (Chatti et al 2012; Papamitsiou & Economides, 2014). Joint enquiry creates informed learning design, allocation of resources, predictive modeling, cultural change, and best practice.

• **Develop infrastructure.** Universities have differences in structure and culture through which technology and learning intersect. Inquiry can explicate strengths and weaknesses across learning management systems, applications, etc. Inquiry may guide university infrastructure for student and faculty support, the jurisdiction of management, training of educational stakeholders, organizational change, and universal design for digital learners.

• **Share information.** In order for LA to grow, investigators must distribute findings across stakeholders, organizations, and professions. SoTL (and SoTEL) is a shared commitment to evaluate what constitutes best educational practice, enabling informed decisions related to educational and organizational goals (Scheffel, Drachsler, Stoyanov & Specht, 2014). Scholarship of knowledge, discovery, and application will inform pedagogy and support capacity.

**Conclusion**

Employing LA through SoTL in distance education may transform teaching and learning. This single case study enabled the investigator to look at connections and their implications for best practice in distance education on a very small scale. There are ample opportunities for the scholarship of discovery, integration, application, and teaching and learning to bridge educational and technical spheres.

**Key Messages**

• Low-cost and open source LA tools may improve teaching and learning with instructors for whom LA is novel.

• Instructors may use LA to gather and discover patterns in the data.

• Instructors may apply LA for individuals, cohorts, and programs.

**References**


