The purpose of the Online Learning Consortium (OLC) is to help learning organizations continually improve the quality, scale, and breadth of online programs according to their own distinctive missions, so that education will become a part of everyday life, accessible and affordable for anyone, anywhere, at any time, in a wide variety of disciplines.

To learn more about submission guidelines for authors and subscription information, please visit: http://onlinelearningconsortium.org/publications/olj_guidelines.

Publications assistants:

Crossfield Professional Writers and Damien Bilka : Copyeditors

Beth Meigs: Layout Editor

Printed in the United States of America 0 9 8 7 6 5 4 3 2 1

ISSN: 2472-5749 (print)

ISSN: 2472-5730 (online)

The Online Learning Consortium, Inc. is a consortium of higher-education providers sharing the common bonds of understanding, supporting and delivering education via asynchronous learning networks (ALNs). With the mission of providing learning to anyone anywhere, OLC seeks to provide new levels of learning capability to people seeking higher and continuing education. For more information about OLC please visit: http://onlinelearningconsortium.org.

Online Learning

Volume 20 Issue 2 – June 2016

SPECIAL ISSUE: LEARNING ANALYTICS

Table of Contents

SECTION I: LEARNING ANALYTICS

Introduction:
Analytics Research and Case Studies in Online Learning

Karen Vignare, Ph.D.
KV Consulting

Patsy Moskal, Ed.D.
University of Central Florida

We are pleased to present this Online Learning Journal special issue devoted to learning analytics for online teaching and learning. The nine papers contained here provide a range of information including reviewing the literature, examining frameworks in development, presenting a large scale analysis on the effectiveness of learning modalities from the PAR Framework, examining an international study of learning, and providing real-world learning analytics case studies on transfer, facilitation, and medical education. Each of these papers provides new and informative research that we hope can help readers make decisions about applying analytics within the context of their own online teaching and learning environments.

While research studies on analytics are beginning to populate journals and conferences, many of those articles are aimed at a more limited audience of researchers. This collection of articles presents readers with information about designing environments within online learning while also highlighting studies that expand upon what exists in currently published research. The authors here represent a significant contribution to practical decision making for administrators, insights for faculty teaching online courses, and works for other researchers to build upon.

Preview of Articles

In the first article, "Learning Analytics Methods, Benefits, and Challenges in Higher Education: A Systematic Literature Review," Avella, Kebritchi, Nunn and Kanai (2016) provide a comprehensive literature review. As the authors note, there is a need for a review of the recent research within the learning analytics field. Given the importance of this topic, it is important that universities and colleges investigate applying learning analytics, and that practitioners understand methodologies as well as benefits and challenges within learning analytics. The opportunities to leverage analytics are indeed powerful, but it is important to understand the multiple methods in the field such as visualization, network analysis, semantic analysis and data mining techniques like predictive, clustering, and relationship mining. Delineated benefits of utilizing learning analytics include improving course offerings, student outcomes, curriculum development and instructor effectiveness. The authors include multiple challenges to the uptake of learning analytics including data quality, analytics used without a connection to learning sciences, and ethical and privacy issues surrounding potentially sensitive data. The systemic literature review is an excellent jumping off point for the special issue.

The next two papers from Australian authors explore frameworks which support the use of analytics within learning environments. Readers will expand their knowledge and perspectives given the thoughtful approach of these authors. The first framework focuses on applying analytics with a primary focus on student retention. The West, Heath and Huijser (2016) article entitled, “Let’s Talk Learning Analytics: A Framework for Implementation in Relation to Student Retention,” presents a dialogical tool for the advancement of learning analytics implementation for student retention. The authors were supported by the Australian Government’s Office for Learning and Teaching and include perspectives from Australian higher education institutions. The project took a mixed-method approach including a survey with both the institutional level and teaching and other academics staff level with an interest in student retention. The paper includes a series of interviews. The framework was then shared with other colleagues to continue the refinement process. The authors hope their tool will be used to inform discussion about how learning analytics might be implemented when an institution is focused on student retention.

The second framework integrates two widely disseminated best practices approaches for designing learning and aims to create an instrument that would inform practices for online and blended learning courses. Jabar and Albion (2016) present the findings in their paper, “Assessing the Reliability of Merging Chickering & Gamson’s Seven Principles for Good Practice with Merrill’s Different Levels of Instructional Strategy (DLISt7).” The study was conducted across five faculties at a regional Australian multi-campus university. The authors used an exploratory and confirmatory factor analysis to validate Merrill’s Different Levels of Instructional Strategy, leveraging the data that had been collected.

Authors James, Swan and Daston (2016) provide a review of a large dataset compiled within the Predictive Analytics Reporting (PAR) Framework by comparing students taking only onground courses, students taking only online courses, and students taking a mixture of both at five primarily onground community colleges, five primarily onground four-year universities, and four primarily online institutions. This exciting work entitled, “Retention, Progression and the Taking of Online Courses,” provides further proof that online courses can provide both flexibility and access while improving student completion.

The results suggest that taking online courses is not necessarily harmful to students’ chances of being retained. While the PAR Framework dataset represents a microcosm of institutions across American universities, it does include a more representative sample of institutions serving nontraditional students. It is clear from other work including IPEDS recent reports that these students are taking more accessible course modalities like online and blended courses. Their research also reveals that essentially no difference in retention between delivery mode was found for students enrolled in primarily onground four-year universities participating in the PAR Framework. At participating primarily online institutions, students blending their courses had slightly better odds of being retained than students taking exclusively onground or exclusively online courses. This report furthers other seminal research that reviews retention in online and blended learning.

The next paper explores the less structured online learning environment of an interactive social network across multiple countries to examine whether the application of learning analytics could be informative in such a network. “The Assessment of Learning in Digital Interactive Social Networks: A Learning Analytics Approach,” by Wilson, Scalise and Gochyvev (2016) examined the learning data within the Assessment and Teaching of 21st Century Skills (ATC21S) project on the “ICT Literacy — Learning in digital networks.” This project, sponsored by Cisco, Intel and Microsoft, helps educators around the world enable students with the skills necessary for success in their future career and college goals. The instrument developed models of learning progression. Results indicate that the new constructs and the new item forms and analytics that were employed, can be used in a large-scale digital environment.

Nadasen and List (2016), in their piece entitled, “Using Community College prior Academic Performance to Predict Re-enrollment at a Four-Year Online University,” provide readers with a case study from a large public online university where analytics were used to predict student success of transfer students. The research study examines student learner characteristics, course-taking behaviors from prior community colleges attended, and first-term GPA at a four-year institution to predict the likelihood of re-enrollment for 8,200 students. The logistic regression models showed that gender, age, and first term GPA at the four-year institution were significant predictors of re-enrollment. Researchers and practitioners will have a greater understanding of how community college factors influence the progression and success for transfer students at four-year institutions.

The authors Phirangee, Demmans Epp and Hewitt (2016) compare instructor facilitation to peer facilitation within online courses. The paper titled, “Exploring the Relationships Between Facilitation Methods, Students’ Sense of Community, and Their Online Behaviors,” compared the two facilitation methods and discovered that students participated more in instructor-facilitated online courses where they wrote, edited and reread posts more, and created more connections to others’ writing, than students in peer-facilitated online courses. They identified student activity patterns and described differences in how those patterns manifest themselves. Their findings also show that instructor-facilitated online courses had a stronger sense of community than peer-facilitated online courses.

Finally, the last two articles explore analytics as a lever to improve student learning. The first article by Wise, Vytasek, Hausknecht, and Zhao (2016) focuses on students as users of analytics. The paper, “Developing Learning Analytics Design Knowledge in the “Middle Space” The Student Tuning Model and Align Design Framework for Learning Analytics Use,” pulls from the literature challenges to using student-facing learning analytics. It also explores a process model for understanding students’ use of learning analytics as part of self-regulation theory. As captured in the title, the authors’ exploratory work refers that this approach will need to be regularly tuned for students to leverage analytics. The resulting framework pulls together constructs and informs how pedagogical design can be improved to leverage analytics. The authors note that this initial inquiry will require more research.

The final article in the issue shows how learning analytics can improve student learning in a medical environment. Using large amounts of data from a single virtual patient case, Poitras, Naismith, Doleck and Lajoie (2016) used subgroup discovery methods to identify student learning patterns. The paper, “Using Learning Analytics to Identify Medical Student Misconceptions in an Online Virtual Patient Environment,” describes how the researchers examined the data to find patterns of learning. The analysis was then used to create an adaptive algorithm where feedback was received by students leading to an improvement in learning outcomes. Adaptive systems are able to leverage analytics in a way that can be very informative to student learning.

We hope that the collection of articles in this special issue strengthens the research base for leveraging analytics within online learning. Some of the articles inform the use of analytics in a variety of learning modalities. However, online learning modalities generally provide researchers with the ease of collecting much more data to explore, analyze and test hypotheses. The field will need significantly more research as well as researchers exploring both small and large datasets to provide practitioners with informative and effective practices. Authors from these papers help to extend the existing groundwork and indicate the continued need for research within analytics for understanding the impact of online learning by addressing issues such as building frameworks to inform design and improve retention, building a research base for student use of analytics, examining how to build large enough datasets that inform practices within the transfer ecosystems of higher education, and leveraging analytics to design adaptive systems to improve student outcomes.

References

Avella, J. T., Kebritchi, M., Nunn, S.G. & Kanai, T. (2016). Learning analytics methods, benefits, and challenges in higher education: A systematic literature review. *Online Learning,* 20 (2), 13 -29.

Jabar, S. I., & Albion, P. R., (2016). Assessing the reliability of merging Chickering and Gamson's seven principles for good practice with Merrill's different levels of instructional strategy (DLISt7). *Online Learning,* 20 (2), 51 - 74.

James, S., Swan, K. & Datson, S. (2016). Retention, progress and the taking of online courses. *Online Learning,* 20 (2), 75 - 96.

Nadasen, D & List, A. (2016). Using community college prior academic performance to predict re-enrollment at a four-year online university. *Online Learning,* 20 (2), 120 - 133.

Phirangee, K., Demmans Epp, C. & Hewitt, J. (2016). Exploring the relationships between facilitation methods, students' sense of community, and their online behaviors. *Online Learning,* 20 (2), 134 - 154.

Poitras, E.G., Naismith, S. M., Doleck, T. & Lajoie, S. P. (2016). Using learning analytics to identify medical student misconception in an online virtual patient environment. *Online Learning,* 20 (2), 183 - 194.

West, D., Heath, D. & Huijser, H. (2016). Let's talk learning analytics: A framework for implement in relation to student retention. *Online Learning,* 20 (2), 30 - 50.

Wilson, M., Scalise, K. & Gochyvev, P. (2016). The assessment of learning in digital interaction social networks: A learning analytics approach. *Online Learning,* 20 (2), 97 - 119.

Wise, A. F., Vytasek, J. M., Hausknecht, S. & Zhao, Y. (2016). Developing learning analytics design knowledge in the 'Middle Space': The student tuning model. *Online Learning,* 20 (2), 155 - 182.

Learning Analytics Methods, Benefits, and Challenges in Higher Education: A Systematic Literature Review

John T. Avella, Mansureh Kebritchi, Sandra G. Nunn, Therese Kanai
University of Phoenix

Abstract

Higher education for the 21st century continues to promote discoveries in the field through learning analytics (LA). The problem is that the rapid embrace of of LA diverts educators' attention from clearly identifying requirements and implications of using LA in higher education. LA is a promising emerging field, yet higher education stakeholders need to become further familiar with issues related to the use of LA in higher education. Few studies have synthesized previous studies to provide an overview of LA issues in higher education. To address the problem, a systemic literature review was conducted to provide an overview of methods, benefits, and challenges of using LA in higher education. The literature review revealed that LA uses various methods including visual data analysis techniques, social network analysis, semantic, and educational data mining including prediction, clustering, relationship mining, discovery with models, and separation of data for human judgment to analyze data. The benefits include targeted course offerings, curriculum development, student learning outcomes, behavior and process, personalized learning, improved instructor performance, post-educational employment opportunities, and enhanced research in the field of education. Challenges include issues related to data tracking, collection, evaluation, analysis; lack of connection to learning sciences; optimizing learning environments, and ethical and privacy issues. Such a comprehensive overview provides an integrative report for faculty, course developers, and administrators about methods, benefits, and challenges of LA so that they may apply LA more effectively to improve teaching and learning in higher education.

Introduction

The advancement of technology has provided the opportunity to track and store students' learning activities as big data sets within online environments. Big data refers to the capability of storing large quantities of data over an extended period and down to particular transactions (Picciano, 2012). Users can take big data from different sources to include learning management systems (e.g., Blackboard), open source platforms (e.g., Moodle), open social platforms (e.g., LinkedIn), and different web tools such Meerkat-Ed and Snapp (Reyes, 2015). Similar to decision making driven by data, analytics refers to the scientific process that examines data to formulate conclusions and to present paths to make decisions (Picciano, 2012). According to Brown (2012), the process of systematically collecting and analyzing large data sets from online sources for the purpose of improving learning processes is called learning analytics (LA). LA is an emerging field in education. Experts in online learning in American higher education predict that within the next few years learning analytics will be widely used in online education to identify students' pattern of behaviors and to improve students' learning and retention rates.

Learning analytics, educational data mining, and academic analytics are closely related concepts (Bienkowski, Feng, & Means, 2012; Elias, 2011). Educational data mining focuses on developing and implementing methods with a goal of promoting discoveries from data in educational settings. It examines patterns in a large data set related to students' actions. The methods may be utilized to form a better understanding of the educational settings and learners. Hung, Hsu, and Rice (2012) defined data mining as data analysis techniques which when applied extract hidden knowledge consisting of tasks consisting of *pattern discovery* as well as *predictive modeling*. Romero and Ventura (2010) provided a definition of educational data mining that uses data mining algorithms with the objective of solving educational issues. Academic analytics refers to an application of the principles and tools of business intelligence to academia with the goal of improving educational institutions' decision-making and performance (Campbell, De Blois, & Oblinger, 2007). Academic analytics combines "large data sets, statistical techniques, and predictive modeling" (Campbell et al., 2007, p. 42).

Learning analytics uses predictive models that provide actionable information. It is a multi-disciplinary approach based on data processing, technology-learning enhancement, educational data mining, and visualization (Scheffel, Drachsler, Stoyanov, & Specht, 2014). The purpose of LA is to tailor educational opportunities to the individual learner's need and ability through actions such as intervening with students at risk or providing feedback and instructional content. Conversely, educational data mining tries to generate systematic and automated responses to learners. While LA focuses on the application of known methods and models to address issues affecting student learning and the organizational learning system, educational data mining focuses on the development of new computational data analysis methods (Bienkowski et al., 2012).

There has been some criticism that higher education managers and the economic framing of education drive the process of big data mining (Clow, 2013); however, empirical studies indicated that LA can be useful for improving education. LA increases awareness of learners and educators in their current situations that can help them make constructive decisions and more effectively perform their tasks (Scheffel et al., 2014). One of the main applications of learning analytics is tracking and predicting learners' performance as well as identifying potential problematic issues and students at risk (EDUCAUSE, 2010; Johnson, Smith, Willis, Levine, & Haywood, 2011). Some universities have already used LA in various courses to improve learning. For example, Purdue University used predictive modeling based on data collected from the course management system to identify students at risk and provide intervention. The University of Alabama improved student retention by forming a predictive model for students at risk based on the large data set of learners' demographics. In another case, Northern Arizona University connected resource use, risk level, and students' achievement by forming a predicting model to identify which students would benefit from which resource (Campbell et al., 2007). These are some examples of pioneer higher education institutions that applied LA.

Although there have been studies related to using LA in higher education institutions within the last several years, LA is still an emerging field of education. Higher education stakeholders including leaders, administrators, instructors, and course developers need to become familiar with LA methods and application in higher education (Scheffel et al., 2014). The problem is few studies have synthesized the previously conducted studies or provided a combined overview of issues concerning the use of LA in higher education. To address this literature gap and enhance application of LA in higher education, this study conducted a literature review. It provides an overview of methods, benefits, and challenges of using LA in higher education institutions for administrators, instructors, and course developers who are not expert in LA and need to develop a basic understanding about LA. As the use of LA is becoming increasingly popular and urgent in higher education, providing such an overview is critical to enhancing higher education stakeholders' understanding about LA.

Method

To address the research problem, researchers conducted a literature review using the procedure suggested by Cooper (1988) for synthesizing the literature. This systematic procedure helped to (a) formulate the problem, (b) collect data, (c) evaluate the appropriateness of the data, (d) analyze and interpret relevant data, and (e) organize and present the results. Then results were compared with current issues in a large higher education institution.

Formulating the problem. The problem is that embracing LA in evaluating data in higher education diverts educators' attention from clearly identifying methods, benefits, and challenges of using LA in higher education. These three key components need further clarification for higher education stakeholders to help them effectively apply learning analytics in higher education. Educators have to go through the daunting task of sifting through the literature to become familiar with LA methods, benefits, and challenges. To help solve the problem, the following questions guided this review:

1. What are the methods for conducting learning analytics in education?
2. What are the benefits of using learning analytics in education?
3. What are the challenges of using learning analytics in education?

Further identifying and describing LA methods, benefits, and challenges can better help educators in higher education to incorporate LA to improve students' learning.

Data collection. The purpose of data collection was to find empirical studies including quantitative, qualitative, mixed methods, and literature reviews published in peer-reviewed journals since 2000 to identify methods, challenges, and benefits of LA in higher education. The keywords that were used included *learning analytics and methods*, *learning analytics and benefits*, and *learning analytics and challenges*. Other keywords included *data mining and education*, *learning analytics and education*, and *learning analytics*. The databases used for literature research included Google Scholar, Educational Resources Information Center (ERIC), ProQuest, and EBSCO HOST.

Data evaluation and analysis. Based on the described procedure, 112 articles were found. Of these, 10 focused on issues related to learning analytics methods, 16 on benefits, and 18 focused on challenges. The remaining articles were excluded from this review because they could not be used to address the main three questions of the study. Only articles that were directly related to LA methods, benefits, and challenges and helped answer the three research questions were included in this review. The method described by Cooper (1988) was appropriate to guide a systematic review of the literature. The researchers exhausted the literature using the above-described procedure, keywords, and databases. Further, researchers limited the search of the literature to the specified keywords and databases. Therefore, this literature may not include sources not available via the searched criteria and databases. Table 1 provides the citations of sources included in the results section.

Table 1 *Sources Found Corresponding to the Research Questions*

Focus	Sources
Learning Analytics Methods and Analyses	Baker (2010) Baker and Yacef (2009) Bienkowski, Feng, and Means (2012) Campbell and Oblinger (2007) Clow (2012) Clow (2013) Dawson and Siemens (2014) Johnson, Levine, Smith, and Stone (2010) Reyes (2015) Romero and Ventura (2010)
Learning Analytics Benefits	Althubaiti and Alkhazim (2014) AlShammari, Aldhafiri, and Al-Shammari (2013) Arnold and Pistilli (2012 Armayor and Leonard (2010) Bhardwaj and Pal (2011) DiCerbo (2014) Dietz-Uhler and Hurn (2013) Grummon (2009) Hsinchun, Chiang, and Storey (2012) Hung and Zhang (2012) Jantawan and Tsai (2013) Kostoglou, Vassilakopoulos, and Koilias (2013) Mardikyan and Badur (2011) Picciano (2012) Sharda, Adomako Asamoah, and Ponna (2013) Xu and Recker (2012)
Learning Analytics Challenges	Bottles, Begoli, and Worley (2014) Brown (2012) Buckingham Shum and Ferguson (2012) Dyckhoff, Zielke, Bültmann, Chatti, and Schroeder (2012) Johnson, Smith, Willis, Levine, and Haywood (2011) Kay, Korn, and Oppenheimer (2012) Ferguson (2012) Fournier, Kop, and Sitlia (2011) Lias and Elias (2011) McNeely and Hahm (2014) Pea (2014) Picciano (2014) Schroeder (2012)

Sclater (2014a)
Sclater (2014b)
Slade and Prinsloo (2013)
Vahdat et al. (2015)
West (2012)

Results

Based upon the literature review, the results obtained to answer the three research questions are provided in this section.

Learning Analytics Methods in Education

The following methods and analysis approaches inform faculty, educators, and administrators in higher education who are not experts in LA about the available methods reported in the literature. Such an overview provides an integrative report for educators and saves them from the daunting task of a literature search to become familiar with different LA methods.

Learning analytics process. With the current advent of both blended and online learning opportunities, big data and learning analytics are predicted to play a significant role in education in future years. When discussing learning analytics methods in education, it is important to provide a background regarding the flow of analytical information. The flow of analytical information can be traced from the students to the stakeholders within the framework of a hierarchy. When provided the opportunity to offer input and make recommendations, stakeholders can help enrich the learning experiences of students (Reyes, 2015). Researchers play a role as they validate and report their research results to inform stakeholders of best practices. Further, learning analytics also provides insight to instructors and students in the educational setting.

To streamline the flow of information and provide a structured process for collecting and analyzing data in learning analytics, researchers suggested a macro-level process for conducting learning analytics in educational settings. Campbell and Oblinger (2007) proposed five stages of capturing data, reporting the data pattern and trends, predicting a model based on the data by using statistical regression, acting by using an intervention based on the model to improve learning, and refining the developed model. Similarly, a learning analytics cycle was suggested by Clow (2012, 2013) in which researchers collect data from the learners, process the data into metrics, and use the results to perform an intervention that affects the students. The cycle continues as researchers collect additional data from the students for the next cycle of learning analytics.

Learning analytics analysis. Learning analytics focuses on data related to learners' interactions with course content, other students, and instructors. LA integrates and uses analysis techniques related to data mining, data visualization, machine learning, learning sciences, psychology, social network analysis, semantics, artificial intelligence, e-learning, and social aspects (Bienkowski et al., 2012; Dawson & Siemens, 2014). Social network analysis includes analysis of relationships between learners as well as between learners and instructors to identify disconnected students or influencers. Social analysis refers to the analysis of metadata to determine learners' types of engagement within educational settings (Bienkowski et al., 2012).

Data visualization tools and techniques. Visual data analysis includes highly advanced computational methods and graphics to expose patterns and trends in large, complex datasets (Johnson, Levine, Smith, & Stone, 2010). One of the standard techniques is visual interactive principal components analysis; it can be used to reduce many variables into a few by finding elements within datasets. Table 2 shows websites that offer tools for data visualization. Gapminder uses the visual interactive approach to help analyze datasets. IBM Many Eyes has tools such as map-based clouds, charts, and graphs to create a visualization. FlowingData allows users to upload their data and create visualizations. A variety of additional visual analysis websites and tools are gathered by the National Visualization and Analytics Center and available at the Visualization Community website (Bienkowski et al., 2012).

Table 2 *Data Visualization Websites and Tools*

Website and Application	URL
Gapminder: reduces datasets into few	http://www.gapminder.org
IBM Many Eyes: creates data visualization	http://www.boostlabs.com/ibms-many-eyes-online-data-visualization-tool/
FlowingData: uploads the data and creates visualization	https://flowingdata.com/
Visualization Community: includes data visualization tools and websites	http://vacommunity.org/

Additionally, learning analytics uses educational data mining methods to analyze large datasets. Within educational data mining, researchers currently use a variety of popular methods. Classified into five categories, these methods consist of *prediction*, *clustering*, *relationship mining*, *discovery with models*, and *separation of data for use in the process of human judgment* (Baker, 2010; Baker & Yacef, 2009; Romero & Ventura, 2010). The final two categories are of significance within the field of education. This section discusses each of the five categories in detail below.

Predication. *Predication* involves developing a model that uses both a predicted variable and predicator variables. A predicted variable represents a particular component of the data, whereas predicator variables consist of a combination of other data elements. Researchers classify predication into three categories known as classification, regression, and density estimation. Baker (2010) described the three categories as classification methods with the use of decision trees, logistic regression, and support vector machine regression. Regression centers around a continuous variable as the predicted variable. Further, it uses linear regression, neural networks, and support vector machine regression. For density estimation, a probability density function is the predicted variable and the use of kernel functions.

Clustering. *Clustering* entails the discovery of a set of data points that form a logical group together. Therefore, observation reveals the resultant formation of some clusters from the full dataset. The use of clustering becomes most valuable when the categories within a group are unknown. How appropriate the set of clusters is may be evaluated by how well the set of clusters fits the data. Baker (2010) asserted that the goal of clustering involved the discovery of data points that formed a natural group together as well as the full dataset. By dividing a collection of data into logical clusters, researchers can assess how cluster sets explain the meaning of the data.

Relationship mining. The method of *relationship mining* focuses on the goal of discovering relationships between variables in a set comprised of a large number of variables. Forms of relationship mining may include learning which variables are related to a single variable or discovering what is the strongest relationship between two variables. Two criteria are necessary for relationship mining: statistical significance and interestingness (Baker, 2010).

Discovery with models. In the next method known as *discovery with models*, the goal is to develop a model using one of the following methods: predication, clustering, or knowledge engineering. Knowledge engineering uses human reasoning for model development. When using the discovery with models method, a prediction model influences a model's generalization across different contexts (Baker, 2010).

Separation of data for use in the process of human judgment. Researchers classify the *separation of data for use in the process of human judgment* method as a visualization method, in which educational data have a particular structure and meaning rooted within that structure. This method possesses two distinct goals identification and classification. Baker (2010) cited the importance of distilling data for identification when the display of data permits easy identification of well-known patterns which may be difficult to express formally. The learning curve represents an example of this concept. For example, the x-axis represents opportunities to practice a particular skill while the y-axis represents performance. This graphical representation can display Performance as the percentage correct or the amount of time that it takes to respond.

Learning Analytics Benefits in Education

Examination of the literature reveals how the use of big data is beneficial for higher education and includes various aspects from learning analytics that closely examine the educational process to improve learning. Another benefit includes the use of academic analytics that make alterations as a result of the application of algorithms to various points of data to improve learning. Through careful analysis of big data, researchers can determine useful information that can benefit educational institutions, students, instructors, and researchers in various ways. These stakeholder benefits include targeted course offerings, curriculum development, student learning outcomes and behavior, personalized learning, improved instructor performance, post-educational employment opportunities, and improved research in the field of education.

Identifying target courses. An initial benefit that evolves from using big data analysis in education is the ability of educational institutions to identify targeted courses that more closely align with student needs and preferences for their program of study. By examining trends in student enrollment and interests in various disciplines, institutions can focus educational and teaching resources in programs that maximize student enrollment in the most needed areas of study. Schools can better predict graduate numbers for long-term planning of enrollment (Althubaiti & Alkhazim, 2014).

Curriculum improvement. Using big data allows instructors to make changes and adjustments to improve curriculum development in the educational system, such as in the use of curricular mapping of data (Armayor & Leonard, 2010). Through the analysis of big data, educators can determine weaknesses in student learning and comprehension to determine whether or not improvements to the curriculum may prove necessary. Instructors can engage in educational strategic planning to ensure that the learning curriculum targets student needs to maximize learning potential.

Student learning outcome, behavior, and process. Another key benefit of big data and text mining focuses on the ability of schools and instructors to determine student learning outcomes in the educational process as well as determine how to improve student performance (Bhardwaj & Pal, 2011). Researchers noted that the use of educational data mining contributed to positive results in the learning process (AlShammari, Aldhafiri, & Al-Shammari, 2013). Analysis of the data can help educators understand the student learning experience through learner interactions with technology tools such as e-learning and mobile learning (Hung & Zhang, 2012). Use of big data also reveals learning behavior, the impact on adaptive learning, and level of persistence (DiCerbo, 2014) in the learning process. By understanding the effects on learner outcomes, use of this data also reveals how to make improvements in student learning and performance in academic coursework. Therefore, LA allows instructors to evaluate forms of knowledge and adjust educational content accordingly.

Personalized learning. Arnold and Pistilli (2012) discussed an early intervention system that demonstrates the benefits and power of learning analytics. As an example, Course Signal provides students with real-time feedback. The components of students' grades, demographic characteristics, academic background, and demonstrated effort are all addressed. The system employs a personalized email and a stoplight, specific color method to indicate progress or lack thereof. Using learning analytics, the concept of personalized learning reveals student success. Dietz-Uhler and Hurn (2013) asserted that course designers do not account for students who do not begin specific coursework at the same learning stage and who do not proceed, learn, and master course competencies at the same pace. Learning analytics allows faculty to use data collected by the learning management system to observe the frequency of student login. Instructors can also see student interaction within the course, total engagement, pace, and grades. These components serve as predictors of students' potential success or failure. Learning analytics allows for real-time reception of the pertinent data, review as well as the incorporation of data, and real-time feedback for every student.

Improved instructor performance. Using this data also helps to assess instructor performance (Mardikyan & Badur, 2011). The use of data provides an opportunity to improve instructor development so that instructors are better prepared to work with students in a technological learning environment. Through the acquisition of data generated from instructor usage of technology and research tools in online libraries (Xu & Recker, 2012), analysts can determine online behaviors by educators. Therefore, use of this information can help identify areas in need of improvement by the instructor to facilitate enhanced instructor-student interactions in the educational environment.

Post-educational employment. Using big data allows educational institutions to identify post-education employment opportunities for graduates and help target education that more closely aligns with employment market needs. It can also predict graduate employment, unemployment, or undetermined situations about job opportunities (Jantawan & Tsai, 2013). Using big data can help stakeholders in the educational system better understand vocational prospects for students and better assess student learning programs for occupational compatibility (Kostoglou, Vassilakopoulos, & Koilias, 2013). In a global learning environment, this type of information not only can facilitate better educational and post-education vocational planning, but also may prove useful to organizations as they make hiring and budgeting decisions for college graduates in different disciplines.

Learning analytics practitioners and research community. The research community also benefits from the use of big data in education. Researchers can more easily share information and collaborate. They can identify gaps between industry and academia so that research can determine how to overcome problems. Also, useful data analysis represents an important component of the ability of scholars to generate knowledge as well as continue to progress in research disciplines (Sharda, Adomako, Asamoah, & Ponna, 2013). However, these benefits are also offset by the need for trained personnel who can use and apply analytics appropriately. Current researchers note a looming future gap in practitioners possessing requisite analytical skill sets in the area of business intelligence and analytics. Picciano (2012) noted a lack of sufficiently trained database administrators and designers to address present needs. This issue has become an important focus for academic and business researchers seeking to overcome this problem through improved education in this area (Hsinchun, Chiang, & Storey, 2012).

Review of the benefits illustrates the usefulness of big data in education. Data analysis provides educational stakeholders a comprehensive overview of the performance of the institution, curriculum, instructors, students, and post-educational employment outlooks. It also provides scholars and researchers with needed information to identify gaps between education and industry so that educators and institutions can overcome these deficiencies in course offerings. More important, the ability of big data to provide these revelations can help the field of education make significant progress to improve learning processes. Further, LA promotes education as it strives to keep pace with the growth of information systems and new technologies. Therefore, the use of LA can help stakeholders understand how well the higher education process is working (Grummon, 2009). The use of learning analytics can prove useful in higher education performance evaluation at the university level. It can provide higher education institutions, instructors, and students improved metrics by which to measure the effectiveness of teaching methods, the engagement of learners in the educational environment, and the efficiency of the learning process using technology.

Learning Analytics Challenges in Education

The review of the literature revealed the LA challenges about data tracking, data collection, data analysis, a connection with learning sciences, learning environment optimization, emerging technology, and ethical concerns regarding legal and privacy issues.

Data tracking. The digital tracking of information is a technique used by analysts to determine how best to present new learning opportunities as the wave of education continues to move forward into the second decade of the 21st Century. The tracking of big data represents the monitoring system. Current trend tracking indicators regarding the delivery and dissemination of instruction depend on the learning management system used by the institution. Platforms such as Moodle, Canvas, EPIC, and Blackboard have the capability to track the number of times an individual logs into the course room. These platforms also provide significant documentation to determine how involved the student was upon their login. Such tracking provides those who plan and implement new educational programs with valuable information. The monitoring reveals how engaging the curriculum presented is, as well as identifying areas that cause confusion (Brown, 2012).

Data collection. The collection of data can be a challenge when looking at LA. Nonetheless, it represents an important component in planning for continued implementation of educational program growth (Bottles, Begoli, & Worley, 2014). Educators must consider several elements. They must consider the availability of resources at a venue. Next, instructors must establish a viable social platform as it directly relates to interactions between learners to synthesize the educational content. Finally, instructors must discriminate whether the learner population possesses the requisite suitability for this type of learning environment and knowledge acquisition. Besides these challenges, gaps exist because of the inability to share proprietary information gathered by the institution. Further, another problem emerges because the creation of the ideal framework to disseminate educational curriculum takes teamwork, especially among the organizations bidding against one another to capture the learner population who want to engage in this type of learning experience.

Evaluation process. An important consideration of data collection concerns how learning analytics has become a force in the evaluation process. As greater amounts of educational resources become available online, there is a subsequent increase in the total data available regarding learning interactions. For learning analytics to help instructor evaluation to function appropriately, data needs to be delivered in a timely and accurate manner (Picciano, 2014). Learning analytics can provide powerful tools for developing meaning from interactions and actions within a higher education learning environment (Fournier, Kop, & Sitlia, 2011). With the unprecedented explosion of available data for online interactions, it is critical for the continued development of the evaluation process. LA can translate from other fields as interest in the data growth in education becomes more focused. Lias and Elias (2011) noted that statistical evaluation of rich data sources already exists within other professions and fields.

Data analysis. Technical challenges also exist from the assimilation of the data analysis because of the presentation format of the data. Erroneous data can skew the findings causing a misinterpretation of the overall population. Such scenarios are commonplace in the online learning environment. For example, an instructor may create a student profile to isolate an assignment that requires grading, test the ease of submission process, or to determine if there are any gaps in the presentation of the curriculum as it appears for students. Creation of a non-existent learner introduces redundant information that appears in the course without identification. This data does not represent student information but rather misinformation created by the instructor that flows into the big data pool of information (McNeely & Hahm, 2014). When manually conducting data analysis, this information can be easily identified from the population. However, working with data collection from the learning analysis vantage point adds a significant margin of error to the outcome of overall results.

Learning sciences connection. According to Pea (2014), personalized learning and learning opportunities demonstrate an inability to leverage learning analytics optimally; therefore, "the endgame is personalized Cyberlearning at scale for everyone on the planet for any knowledge domain" (p. 17). Ferguson (2012) asserted that to optimize and fully understand learning requires understanding how knowledge develops and how to support knowledge development. Further, researchers must understand the components of identity, reputation, and affect. Researchers must find ways to connect "cognition, metacognition, and pedagogy" (Vahdat et al., 2015, p. 299) to help improve learning processes. With a stronger connection to learning sciences, learning analytics can promote effective learning design.

Learning environment optimization. Ferguson (2012) noted that as learners expand the boundaries of the learning management system into open or blended learning settings, researchers must discover the problems faced by students and how to determine success from the learners' perspectives. This process will encumber a shift toward more challenging datasets that may include mobile, biometric, and mood data. Besides the individual learning aspect of learning analytics, researchers are seeking to address another component known as social learning analytics. In this context, social learning analytics focuses on the collaboration and interaction of learners in a socialized learning environment, not just on individual learning outcomes (Buckingham Shum & Ferguson, 2012).

Emerging technology. The full potential of learning analytics relating to learning requires continued and emerging technology that presently remains in the younger stages. This revelation presents a challenge as the technology continues to develop to stay constant with the growth of learning analytics. Further, to fully understand the method and practice of teaching, more research is needed. Research focusing on learning analytics and pedagogy is still in the beginning stages (Dyckhoff, Zielke, Bültmann, Chatti, & Schroeder, 2012).

Ethical and privacy issues. Another issue that emerges about learning analytics concerns the ethical, legal, and risk considerations (Kay, Korn, & Oppenheimer, 2012). Because of dynamic changes in technology as well as how users store data and applications in cloud-based systems, "the challenges of privacy and control continue to affect adoption and deployment" (Johnson et al., 2011, p. 3). Further, the ethical and legal complexities of learning analytics challenge institutions that seek to implement their usage (Sclater, 2014a). For example, these considerations can include obvious areas of privacy considerations such as consent, data accuracy, how to respect privacy, maintaining anonymity, opting out of data gathering, and the potential effects to students. Additional concerns include data interpretation, data ownership, data preservation, sharing data with parties outside of the institution, and proper training of staff members regarding the handling of data (Sclater, 2014b). Further, the question becomes who owns this aggregate data, because having an infrastructure with the capacity to house large amounts of information becomes a daunting task (West, 2012). Because of these different issues, institutions must achieve a balanced approach to safeguard data while also assuring benefits to the educational process through the use of four guiding principles. These principles consist of clear communication, care, proper consent, and complaint (Kay et al., 2012). Institutions must demonstrate adherence to legal and ethical parameters to safeguard student privacy concerns while also achieving the educational goals for students and educators.

To accomplish privacy and educational goals, Slade and Prinsloo (2013) proposed the use of an ethical framework based on six principles that strives to maintain educational effectiveness for instructors and students. The first principle proposes that learning analytics should focus on understanding what works within the educational process and the moral necessity to use information. The second principle advances the notion that students must collaborate with the institution and willingly provide informed consent for institutions to collect, use, and store data. The third principle asserts that a student's identity and education performance is temporary. Therefore, this information should possess an expiration date as well as permit students to request the deletion of data under predetermined guidelines. Principle four notes that student achievement consists of various factors that can demonstrate a diverse range of considerations and complexities. For this reason, data may not always reflect student learning accurately. Principle five states that transparency should serve as an important guideline to ensure that data usage conforms to appropriate standards based on its purpose, who may access the data, and privacy protection. Finally, principle six suggests that institutions cannot afford to ignore the value of learning analytics and its value to learning outcomes. While learning analytics poses different ethical, legal, and risk challenges that stakeholders must understand, proper awareness and scrutiny can ensure that users adhere to ethical and legal parameters to balance privacy protections of students with institutional information needs.

Discussion

Analysis of the literature provided a comprehensive overview of the analytical learning methods, benefits, and challenges regarding the use of big data in education. Examination of these aspects showed positive contributions. However, the literature review also revealed negative considerations regarding the use of learning analytics. For these reasons, a closer examination of these aspects can reveal other insights.

This study revealed that learning analytics is an interdisciplinary field or *Jackdaw* field of inquiry as suggested by Clow (2013). As such, learning analytics selects and uses methods and analysis techniques from other disciplines as appropriate to achieve the goal of improving education. LA offers interventions based on the predicted models grounded in the large datasets in educational institutions. Learning analytics follows a macro-level process of data collection, data analysis, prediction model development, intervention, and refinement of the predicted model (Campbell & Oblinger, 2007; Clow, 2012, 2013). There is a hierarchy framework in which the flow of analytical information traces from students to the stakeholders. This research reveals that the process must include stakeholders to offer input and feedback. Use of an inclusive strategy will lead to an enhanced experience for the students. Learning analytics uses various approaches including visual data analysis techniques, social network analysis, semantic, and educational data mining to analyze the data (Bienkowski et al., 2012; Dawson & Siemens, 2014). In particular, this study described the two widely used approaches of data visualization and education data mining method of Baker (2010), which includes prediction, clustering, relationship mining, discovery with models, and separation of data for human judgment. Learning analytics allows for stakeholders—such as instructors and learners—to receive insights about the educational setting. Further, researchers must validate and report research results as well as promote best practices for stakeholders.

From using these various methods, big data in education exhibits many benefits in the educational environment. Review of the literature revealed how data analytics can yield benefits to education by helping institutions and instructors to improve course offerings. LA can also help establish subject curriculums that more closely align with the needs of students as well as the needs of industry. Further, analytics can provide useful feedback regarding learner outcomes in the educational process as well as the effectiveness of instructors within the classroom. Moreover, analytics can provide useful information to the research community to help identify and resolve gaps between academia and different industries. Though LA appears to demonstrate a comprehensive representation of benefits within education, one must remember that the use of learning analytics is still in the early stages of development. Therefore, additional benefits may emerge as the use of analytics evolves and technology improves. Further, as people acquire greater skills in the field of analytics, more sophisticated uses of big data may also emerge to provide more useful insights in educational environments.

While learning analytics demonstrates benefits to educators and students through improved lectures and student learning outcomes (Scheffel et al., 2014), a close review of the ethical, legal, and risk concerns demonstrate the complexities of balancing challenges with educational concerns. Though current technological capabilities allow the acquisition, storage, and access of data, users must remain current on technology innovations that may provide improvements to these systems and how institutions can balance stakeholder needs with ethical considerations. To accomplish this, people who work with data must possess the requisite training in learning analytics to understand how to use the data productively to achieve meaningful results. Further, users must also understand the various ethical and legal aspects as well as the inherent risks of data usage. To protect privacy and maintain ethical standards, guidelines should inform appropriate data use. Further, accountability at all levels of data use will require an appropriate system of checks and balances. Additionally, data security should be well implemented and maintained to ensure privacy protections. In summary, mechanisms must provide appropriate transparency, data controls by students, information security, and accountability safeguards (Pardo & Siemens, 2014).

Regardless of the insights revealed by these components, the use of big data in education remains a relatively new area of focus in the field of education. Though the methods provide useful information within the current environment, analysts within the educational community must continue to develop and improve the methods needed. By properly refining the use of data to understand its meaning and implications, researchers can determine how best to use the information to improve the educational process at all levels and for all stakeholders.

Recommendations

It is evident that data acquisition using learning analytics demonstrates considerable implications for the future of the educational process. Not only does LA allow institutions to understand trends for planning purposes, but it also allows educators and learners to focus on needed skills for the 21st century. However, users must continue to grow with the acquisition and application of new technology that can provide enhanced capabilities to meet educational needs. To help improve educational outcomes in the learning environment, researchers will need to explore the use of different types of data sets from biometrics, learner moods, and mobile devices (Ferguson, 2012). The use of biometrics, such as fingerprints, facial recognition, iris recognition, and hand or finger geometry, would provide user authentication to verify user access as well as attendance, work, and testing in the educational process (Sayed & Jradi, 2014). Use of this technology would serve as an additional layer of protection for data and would help to mitigate ethical, legal, and risk concerns. Because of the predicted rise in data acquisition and storage, researchers could integrate biometric services into a cloud-based system or different e-learning environments (Peer, Bule, Gros, & Štruc, 2013).

Another recommendation would be to identify ways to better streamline research priorities to ensure the use of the most pertinent information. For example, researchers could focus on refinement of learning analytics to help achieve better learning outcomes. Further, better standardization of assessments could improve student engagement and help establish learner competency levels. The expansion of learning analytics that includes different learning contexts could provide an improved understanding of learner-to-learner interaction within the educational environment (Pea, 2014). By addressing these recommendations, learning analytics will continue to be a valuable means of predicting future trends and needs in the field of education. As online education continues to flourish into the 21st century, educators need to continue to embrace technological changes in the educational sector. By doing so, these technological innovations will continue to serve as the driving force to make online teaching a successful process to achieve learning outcomes for all students.

Contributions, Implications of the Study, and Results

This study has revealed the different methods used by learning analytics not only to show how the use of big data can benefit education but also to reveal the challenges faced by stakeholders in the educational process. The field of learning analytics has been and will continue to expand greatly in part due to the ability to store increased amounts of data and a large diversity of research strands. As a result of greater availability and access to data, learning analytics will provide an increased understanding of the patterns of learner behavior, networks, and interactions. West (2012) cited the example of a total of sixteen states presently utilizing data mining techniques and learning analytics to identify students classified as at-risk. However, this study also revealed that learning analytics underscores the need for greater understanding of how to analyze data to optimize results and use the information to improve the educational process at all levels.

Going forward, schools must recognize the importance of implementing a data-driven approach to education. The use of performance systems allows for increased and more productive decision-making, the identification of trends and problematic areas, and the more efficient allocation of resources. Moreover, performance systems reveal how stakeholders within education must strive to continually improve methods, processes, and knowledge. Doing so ensures that the use of learning analytics continues to serve the needs of educators, students, institutions, and researchers throughout the educational process. Further, using learning analytics will ensure positive steps forward in the educational system as well as ensure greater outcomes for all stakeholders.

Learning analytics is still an emerging field of education. Higher education stakeholders need to become further familiar with learning analytics applications in education (Scheffel et al., 2014). Nonetheless, there is a gap in literature to provide an integrative report about learning analytics in higher education. To bridge the gap, this review of literature provided an overview of the methods, benefits, and challenges regarding the use of learning analytics for leaders, administrators, instructors, and course developers in higher education. It is hoped that the overview enhances the application of learning analytics in higher education and supports the use of large datasets to improve the quality of teaching and learning.

References

AlShammari, I. A., Aldhafiri, M. D., & Al-Shammari, Z. (2013). A meta-analysis of educational data mining on improvements in learning outcomes. *College Student Journal, 47*(2), 326-333.

Althubaiti, A., & Alkhazim, M. (2014). Medical colleges in Saudi Arabia: Can we predict graduate numbers? *Higher Education Studies, 4*(3), 1-8.

Armayor, G.M., & Leonard, S. T. (2010). Graphic strategies for analyzing and interpreting curricular mapping data. *American Journal of Pharmaceutical Education, 74*(5), 1-10.

Arnold, K. E., & Pistilli, M. D. (2012, April 29). Course signals at Purdue: Using learning analytics to increase student success. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 267-270). New York, NY: ACM. doi: 10.1145/2330601.2330666

Baker, R. (2010). Data mining for education. *International Encyclopedia of Education*, *7*, 112-118.

Baker, R., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, *1*(1), 3–16.

Bhardwaj, B. K., & Pal, S. (2011). Data mining: A prediction for performance improvement using classification. *International Journal of Computer Science and Information Security, 9*(4), 136-140.

Bienkowski, M., Feng, M., & Means, B. (2012). *Enhancing teaching and learning through educational data mining and learning analytics: An issue brief*. U.S. Department of Education, Office of Educational Technology. Washington, D.C. Retrieved from http://www.ed.gov/technology.

Bottles, K., Begoli, E., & Worley, B. (2014). Understanding the pros and cons of big data analytics. *Physician Executive*, *40*(4), 6-12.

Brown, M. (2012). Learning analytics: Moving from concept to practice. *EDUCAUSE Learning Initiative*. Retrieved from http://net.educause.edu/ir/library/pdf/ELIB1203.pdf

Buckingham Shum, S., & Ferguson, R. (2012). Social learning analytics. *Educational Technology & Society*, *15*(3), 3-26.

Campbell, J. P., De Blois, P. B., & Oblinger, D. G. (2007). Academic analytics: A new tool for a new era. *Educause Review*, *42*(4), 40-57. Retrieved from http://www.educause.edu/ero/article/academic-analytics-new-tool-new-era

Campbell, J. P., & Oblinger, D. G. (2007). Academic analytics. *Educause*. Retrieved from http://net.educause.edu/ir/library/pdf/pub6101.pdf.

Clow, D. (2012). The learning analytics cycle: Closing the loop effectively. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 134-138). New York, NY: ACM. doi:10.1145/2330601.2330636

Clow, D. (2013). An overview of learning analytics. *Teaching in Higher Education, 18*(6), 683-695. doi:http://dx.doi.org/10.1080/13562517.2013.827653

Cooper, H. (1988). The structure of knowledge synthesis: A taxonomy of literature reviews. *Knowledge in Society, 1*, 104-126.

Dawson, S., & Siemens, G. (2014, September). Analytics to literacies: The development of a learning analytics framework for multiliteracies assessment. *International Review of Research in Open and Distance Learning*, *15*(4), 284-305.

DiCerbo, K. E. (2014). Game-based assessment of persistence. *Journal of Educational Technology & Society, 17*(1), 17-28.

Dietz-Uhler, B., & Hurn, J. E. (2013, Spring). Using learning analytics to predict (and improve) student success: A faculty perspective. *Journal of Interactive Online Learning*, *12*(1), 17-26.

Dyckhoff, A. L., Zielke, D., Bültmann, M., Chatti, M. A., & Schroeder, U. (2012). Design and implementation of a learning analytics toolkit for teachers. *Journal of Educational Technology & Society*, *15*(3), 58-76.

EDUCAUSE. (2010). Next generation learning challenges: Learner analytics premises. *EDUCAUSE Publications*. Retrieved from http://www.educause.edu/Resources/NextGenerationLearningChalleng/215028

Elias, T. (2011). *Learning analytics: Definitions, processes and potential* (Report). Retrieved from http://learninganalytics.net/LearningAnalyticsDefinitionsProcessesPotential.pdf

Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, *4*(5/6), 304–317.

Fournier, H., Kop, R., & Sitlia, H. (2011). The value of learning analytics to networked learning on a personal learning environment. *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge* (pp. 246-250). New York, NY: ACM. doi:10.1145/2567574.2567613

Grummon, P. T. H. (2009). Trends in higher education. *Planning for Higher Education*, *37*(4), 48-57.

Hsinchun, C., Chiang, R. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, *36*(4), 1165-1188.

Hung, J.L., Hsu, Y.C., & Rice, K. (2012). Integrating data mining in program evaluation of k-12 online education. *Educational Technology & Society*, *15*(3), 27-41.

Hung, J., & Zhang, K. (2012). Examining mobile learning trends 2003-2008: A categorical meta-trend analysis using text mining techniques. *Journal of Computing in Higher Education, 24*(1), 1-17. doi:http://dx.doi.org/10.1007/s12528-011-9044-9

Jantawan, B., & Tsai, C. (2013). The application of data mining to build classification model for predicting graduate employment. *International Journal of Computer Science and Information Security, 11*(10), 1-7.

Johnson, L., Levine, A., Smith, R., & Stone, S. (2010). *The horizon report: 2010 edition* (Report). Retrieved from http://www.nmc.org/pdf/2010-Horizon-Report.pdf

Johnson, L., Smith, R., Willis, H., Levine, A., & Haywood, K., (2011). *The horizon report: 2011 edition* (Report). Retrieved from https://net.educause.edu/ir/library/pdf/HR2011.pdf

Kay, D., Korn, N., & Oppenheim, C. (2012). *Legal, risk and ethical aspects of analytics in higher education* (White Paper). Retrieved from http://publications.cetis.ac.uk/wp-content/uploads/2012/11/Legal-Risk-and-Ethical-Aspects-of-Analytics-in-Higher-Education-Vol1-No6.pdf

Kostoglou, V., Vassilakopoulos, M., & Koilias, C. (2013). Higher technological education specialties and graduates' vocational status and prospects. *Education & Training, 55*(6), 520-537. doi:http://dx.doi.org/10.1108/ET-03-2012-0026

Lias, T. E., & Elias, T. (2011). *Learning analytics: The definitions, the processes, and the potential* (Report). Retrieved from http://learninganalytics.net/LearningAnalyticsDefinitionsProcessesPotential.pdf

Mardikyan, S., & Badur, B. (2011). Analyzing teaching performance of instructors using data mining techniques. *Informatics in Education*, *10*(2), 245-257.

McNeely, C. L., & Hahm, J. (2014). The big (data) bang: Policy, prospects, and challenges. *Review of Policy* Research, *31*(4), 304-310.

Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology*, *45*(3), 438-450. doi:10.1111/bjet.12152

Pea, R. (2014). *The learning analytics workgroup: A report on building the field of learning analytics for personalized learning at scale* (Report). Retrieved from https://ed.stanford.edu/sites/default/files/law_report_complete_09-02-2014.pdf

Peer, P., Bule, J., Gros, J. Ž., & Štruc, V. (2013). Building cloud-based biometric services. *Informatica*, *37*(2), 115-122.

Picciano, A.G. (2012). The evolution of big data and learning analytics in American higher education. *Journal of Asynchronous Learning Networks, 16* (3), 9-20.

Picciano, A. G. (2014). Big data and learning analytics in blended learning environments: Benefits and concerns. *International Journal of Artificial Intelligence and Interactive Multimedia*, *2*(7), 35-43.

Reyes, J. A. (2015). The skinny on big data in education: Learning analytics simplified. *TechTrends*, *59*(2), 75-79.

Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, *40*(6), 601-618.

Sayed, M., & Jradi, F. (2014). Biometrics: Effectiveness and applications within the blended learning environment. *Computer Engineering and Intelligent Systems*, *5*(5), 1-9.

Scheffel, M., Drachsler, H., Stoyanov S., & Specht, M. (2014). Quality indicators for learning analytics. *Educational Technology & Society*, *17*(4), 117–132.

Sclater, N. (2014a, September 18). *Code of practice "essential" for learning analytics*. Retrieved from http://analytics.jiscinvolve.org/wp/2014/09/18/code-of-practice-essential-for-learning-analytics/

Sclater, N. (2014b, November). *Code of practice for learning analytics: A literature review of the ethical and legal issues*. Retrieved from http://repository.jisc.ac.uk/5661/1/Learning_Analytics_A-_Literature_Review.pdf

Sharda, R., Adomako Asamoah, D., & Ponna, N. (2013). Research and pedagogy in business analytics: Opportunities and illustrative examples. *Journal of Computing & Information Technology*, *21*(3), 171-183. doi:10.2498/cit.1002194

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, *57*(10), 1510-1529. doi: 10.1177/0002764213479366

Vahdat, M., Ghio, A., Oneto, L., Anguita, D., Funk, M., & Rauterberg, M. (2015). Advances in learning analytics and educational data mining. *Proceedings from 2015 European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, *Belgium*. Retrieved from http://www.idemployee.id.tue.nl/g.w.m.rauterberg/publications/ESANN2015paper1.pdf

West, D. M. (2012, September). Data mining, data analytics, and web dashboards. *Governance Studies at Brookings*, 1-10. Retrieved from Brookings.edu website at: http://www.brookings.edu/~/media/research/files/papers/2012/9/04%20education%20technology%20west/04%20education%20technology%20west.pdf

Xu, B., & Recker, M. (2012). Teaching analytics: A clustering and triangulation study of digital library user data. *Journal of Educational Technology & Society, 15*(3), 103-115.

Let’s Talk Learning Analytics: A Framework for Implementation in Relation to Student Retention

Deborah West and David Heath
Charles Darwin University

Henk Huijser
Xi'an Jiaotong-Liverpool University & Batchelor Institute of Indigenous Tertiary Education

Abstract

This paper presents a dialogical tool for the advancement of learning analytics implementation for student retention in Higher Education institutions. The framework was developed as an outcome of a project commissioned and funded by the Australian Government’s *Office for Learning and Teaching*. The project took a mixed-method approach including a survey at the institutional level (n = 24), a survey of individual teaching staff and other academics with an interest in student retention (n = 353), and a series of interviews (n = 23). Following the collection and analysis of these data an initial version of the framework was developed and presented at a National Forum attended by 148 colleagues from 43 different institutions. Participants at the forum were invited to provide commentary on the usefulness and composition of the framework which was subsequently updated to reflect this feedback. Ultimately, it is envisaged that such a framework might offer institutions an accessible and concise tool to structure and systematize discussion about how learning analytics might be implemented for student retention in their own context.

Introduction

This paper reports on findings from an Australian Government Office for Learning and Teaching commissioned project entitled Learning Analytics: Assisting Universities with Student Retention. One of the primary outputs of this project was a framework based on data collected from institutions and academics, using a mixed method approach featuring two different surveys and a series of interviews, conducted between July, 2014 and February, 2015. Following the surveys and interviews a draft framework of factors relevant to the institutional implementation of learning analytics for student retention purposes was developed. The framework was presented to peers and their feedback elicited at a National Forum held in April 2015, based on which the draft framework was refined to arrive at the version presented here.

The primary purpose of this paper is to present a framework of factors relevant to the institutional implementation of learning analytics, situate it with other thinking in this area, and discuss how the framework might be used in practice. Such a framework may prove a useful tool for institutions engaged in decision making and planning around the use of learning analytics for student retention purposes, particularly because learning analytics initiatives often require people with a wide variety of expertise to work together (e.g. coding and programming, project management, learning and teaching, discipline knowledge, statistics, and potentially a range of others). A secondary purpose is to explain the process of the framework's development and explore the context in which this development occurred. Attention to this second element is critical as it draws attention to the underpinning thinking linked to previous work in the field, and data input highlighting the need for a dialogical tool due to the complexity of learning analytics.

Structurally, the paper begins with a brief literature review in which the two primary concepts associated with the framework are explored: student retention and learning analytics. The literature review also considers other frameworks and tools related to institutional implementation of learning analytics. To conclude the literature review, the research questions and thinking behind these are described followed by the research methods. The results section focuses initially on headline findings from the data, which leads into presentation of the framework. The paper concludes with some propositions about how the framework might be used, discussion of how it fits with other conceptualisations of institutional implementation of learning analytics, and a brief overview of future directions.

Literature Review

Student Retention

Retention, and more broadly student success, are critical but challenging issues within the higher education sector in Australia. Widening participation agendas, which include the desire to have an increased number of students and a broader diversity of students, are a high priority (Bradley, Noonan, Nugent & Scales, 2008). Key cohorts for improving higher education participation and success are Aboriginal and Torres Strait Islander people (Universities Australia, 2013; Behrendt, Larkin, Griew & Kelly, 2012), people from regional areas and people of low socio-economic status (Universities Australia, 2013).

Responding to this challenge, institutions have been working toward adapting their environments and learner support mechanisms. One aspect of this has been the use of technology enabled learning, which can provide increased flexibility and personalised learning opportunities. Despite the efforts made to date, attrition rates continue to be high, particularly for more diverse learner cohorts with a sector average of 17.31% (Koshy, 2014). Additionally, there are concerns around higher attrition rates for students studying online (Dekker, Pechenizkiy & Vleeshouwers, 2009; Rivera & Rice, 2002). There is some speculation around the reasons for this including differences in student demographics, difficulties with effective use of technology and social isolation (Frankola, 2002).

The academic and non-academic factors that can influence retention are complex and varied (Nelson, Clarke, Stoodley & Creagh, 2014). In a study of business students across six Australian universities (n = 7486), Willcoxson and colleagues (2011: 1) reported the following:

> …data strongly indicates that factors related to attrition are generally university-specific and reflect both student characteristics and their responses to the specific institutional culture and environment. The only attrition triggers which span most universities and most years of study are 'lack of a clear reason for being at university' and 'the feeling of having insufficient ability to succeed at university'.

This suggests that attempts to tackle attrition need to be considered and contextualised to specific institutions. It is widely recognised that retention is one facet of student success and that many of the actions that we take to improve outcomes for those students who are 'at risk' can actually benefit all students throughout their learning journey.

Learning Analytics

Linked to technology enabled learning and broader digitalisation of interactions (e.g. libraries, student services engagements) are a range of opportunities to collect and analyse data with the intention of improving the student experience (Gaševic, Dawson, & Siemens, 2015). As Buckingham Shum and Ferguson have observed (2012: 3) "the core proposition is that, as unprecedented amounts of digital data about learners' activities and interests become available, there is significant potential to make better use of this data to improve learning outcomes". While there has been much debate around definitions of learning analytics and attempts to disaggregate learning analytics from educational analytics, educational data mining and academic analytics (Ferguson, 2012; Gaševic et al., 2015; Long & Siemens, 2011; van Barneveld, Arnold, & Campbell, 2012), such debates are not seen as useful in the context of this paper. Learning analytics is an emerging field and most in the sector are more interested in how data can be used rather than the nuances of definitions (Sclater, 2014). For our purposes, learning analytics can be seen as 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" (Long & Siemens, 2011). Sclater (2014) notes that such data can be seen to be on a continuum of use value to different stakeholders (which is more in line with how institutions perceive it) and can be used by these various stakeholders (including students, tutors, educational researchers, unit heads and management) for different purposes, depending on their particular focus.

Mapping Institutional Progress with Learning Analytics

Irrespective of definitions, the implementation of learning analytics in the sector is at an early stage (Colvin, et al., in press; Fiadhi, 2014; Johnson, Adams Becker, Estrada, & Freeman, 2015; Sclater, 2014; West et al., in press). While a variety of case studies in relation to initiatives is available internationally (see Arnold & Pistilli, 2012; Sclater, 2014; Yeado, Haycock, Johnstone, & Chaplot, 2014), these tend to be of a relatively small scale, localised or narrow focus. They are useful to demonstrate what is possible and to share practice, but do not necessarily move others forward on implementation. Efforts to assist in this regard have tended to focus on frameworks for building organisational capacity and/or benchmarking more broadly, and related to technology enabled learning, as well as subsequent application of these to learning analytics. These have included a variety of foci including measuring capability (Marshall and Mitchell, 2002; Marshall, 2006; ACODE, 2014), advocating process (Bichsel, 2012; Campbell and Oblinger, 2007) and/or identifying domains for attention (Arnold, Lonn, & Pistilli, 2014; Arnold, Lynch, et al., 2014; Sclater, 2014). Thus, most frameworks generally incorporate each of these, but it is useful to highlight some of the main elements.

Benchmarking is a critical element of quality control mechanisms and as such the sector often looks for ways to assess our development in key areas. Much work has been done on this element related to technology enabled learning, including the earlier work of Marshall & Mitchell (2002), and subsequently Marshall (2006). Their e-learning Maturity Model (eMM) is intended to assist institutions to deliver e-learning in a sustainable and high quality manner. Building on Chickering and Gamson (1987) they initially presented a model based on levels of capability. However, the second version removed the concept of levels, acknowledging that it implied a hierarchical structure which was unhelpful in a dynamic environment and they moved to one based on dimensions. The overall premise is that there needs to be capability across the dimensions which include: learning; development; support; evaluation and quality control; institutional planning and management. The concept of process around these is overlaid in each dimension. Each dimension then includes a list of statements to prompt reflection and rating. At the core, the model provides a way of assessing how an institution rates in relation to each dimension, from fully adequate to not adequate.

Similar to Marshall (2006), Educause's ECAR model (which is specifically focussed on learning analytics) attempts to provide a framework for measuring capability but places more emphasis on the domains. The domains included in this model are: process, culture, expertise, investment, governance/infrastructure and data/reporting tools. In a similar vein, The Learning Analytics Readiness Instrument (LARI) (Arnold, Lonn, et al., 2014) focusses on 5 domains: ability; data; culture and process; governance and infrastructure; and overall readiness perceptions. It is intended for use by institutions to develop an understanding of both their strengths and areas for development via rating against statements in each domain.

The use of a series of statements to rate one's progress is the hallmark of benchmarking frameworks. The Australian Council on Open and Distance Education (ACODE) outlines and follows this process in the development of their document for benchmarking technology enhanced learning. The document (ACODE, 2014) includes a scoping statement, good practice statement and performance indicators and measures; it also has room for rationale, evidence, and recommendations for eight domains. These domains include governance; planning; information technology (IT) infrastructure; application of technology; staff development; staff support; student training and student support. It is designed as a tool to be used actively by institutions for self-assessment and comparison.

Another version that focussed more specifically on learning analytics, and that incorporates stages and domains, is put forward by Norris and Baer (2013). Their model incorporates the domains of technology infrastructure; processes and practices; culture and behaviour; skills and values; and leadership. It includes three stages of capacity building from getting started to accelerating progress, and lastly to transforming/optimising analytics.

While most frameworks place some emphasis on process, for some this is a central focus; most notably Bichsel (2012) and Campbell and Oblinger (2007). Campbell and Oblinger identify 5 steps, which are: capture; report; predict; act and refine. Bichsel notes that while there is a wide degree of variation in how analytics might be used, it can be conceptualised as having a common cyclical process with the following five stages that loop back as refinement takes place:

1. Identify a strategic question
2. Find or collect relevant data to the question
3. Analyse the data with a view to prediction and insight
4. Form and present the data in understandable and actionable ways
5. Feedback into the process of addressing the strategic question and identifying new ones

Lessons from the Literature

To conclude this brief overview, it is useful to highlight the themes that arise from the available frameworks. They all advocate, to a greater or lesser extent, the need for ongoing refinement and improvement via a process of reflection. They draw attention to the fact that development and progress in learning analytics is a complex, interdependent and dynamic endeavour, which encompasses a variety of stakeholders across the institution. Related to this, a variety of domains need to be attended to, including those most commonly noted: technical infrastructure, policy and governance, skills; support and culture. Implied within these frameworks, and explicit in the retention literature, is the need to undertake development in a way that is appropriate to any given institution. So while it is critical to share examples and strategies, there is a need to work more effectively within an institution to make clear progress. Working from these key ideas, our project sought to develop a framework for moving forward on the use of learning analytics for student retention.

Project Context and Research Questions

As learning analytics is at an early stage and considerable variety exists amongst analytics methods and products, and the individual contexts in which learning analytics might be deployed, it was difficult to hypothesise which themes would emerge most strongly over the course of the study. The project team was aware through its own professional networks and involvement in various forums that many universities were actively exploring and testing different tools, but that widespread and mature usage seemed fairly limited. Consequently, the project adopted a broad and exploratory scope with the intention to elicit information from a broad range of contexts. This is reflected by the two main research questions:

1. What factors are relevant or need to be considered where the implementation of learning analytics for student retention purposes is concerned?
2. How do these factors impact on the implementation of learning analytics for student retention purposes?

Method

Given that this paper is primarily focused on introducing a framework, it is useful to think of the method as having two distinct phases. The first is the initial mixed-method (survey and interview) data collection that informed the development of a draft framework. The second phase represents the National Forum where the framework was presented to peers for feedback and a process of refinement and adjustment that followed.

Phase 1

Institution level survey. The institution level survey, which can be viewed here, was conducted in July and August 2014, and was directed toward senior academic leaders who could give an overview of the institutional strategy around learning analytics. The survey was intended to help build a picture of sector progress with learning analytics more broadly and in relation to student retention, as well as data infrastructure, human resources, policy development and other elements that relate to learning analytics as identified through the literature. It was distributed to Deputy Vice Chancellors (Academic) (DVCA's) at each Australian university via email and promoted via the Universities Australia DVCA forum. Twenty-two Australian institutions participated. The sample was extended to New Zealand (NZ) following a request from the project reference group which resulted in two NZ institutions participating. The survey was built and hosted using the online Qualtrics application and was set to allow anonymous responses.

SPSS was used for the quantitative analysis to produce descriptive statistics while manual coding was used to analyse qualitative questions.

Academic level survey. The academic level survey, which can be viewed here, conducted between September and November, 2014, was targeted at academic staff (e.g. teachers, student support, academic developers etc.), and focused on how they were thinking about and/or using learning analytics in the context of student retention. The survey employed a purposive, snowball sampling strategy to recruit self-selecting individuals (n = 353), whose responses were anonymous, with invitations circulated via existing networks, professional associations and conferences. Analysis involved a range of techniques including frequencies, means testing and tests of significance as appropriate to the individual quantitative questions and manual coding of themes for qualitative questions.

Table 1: Frequency distributions of demographic data about academic level survey participants

Variable	**Category**	**Absolute Frequency**	**Relative Frequency**
Location (n = 351)	Australia	341	97%
	New Zealand	10	3%
Primary Work Role (n = 353)	Teaching Students	188	53%
	Learning Support	47	13%
	Management/Administration	37	11%
	Other	32	9%
	Research	24	7%
	Academic Development	18	5%
	Student Support	7	2%
LMS at Institution (n = 353)	Blackboard	203	58%
	Moodle	124	35%
	Brightspace (D2L)	13	4%
	Sakai	2	1%
	Other	11	3%
Academic Level (n = 351)	Associate Lecturer/Tutor	30	9%
	Lecturer	124	35%
	Senior Lecturer	88	25%
	Associate Professor	30	9%
	Professor	20	6%
	Other	59	17%
Length of employment in Higher Education Sector (n = 345)	Less than 1.5 years	11	3%
	1.5 – 5 years	42	12%
	5 – 10 years	85	25%
	10- 20 years	130	38%
	More than 20 years	77	22%
Involvement in teaching students (n = 353)	Teaches students	276	78%
	Does not teach students	77	22%

Table 1 presents a summary of sample demographics in the form of frequency distributions for a range of employment related indicators. While the sample represents a small proportion of the overall population of Australian academic staff potentially involved with student retention, it produced strong heterogeneity as far as teaching modalities, experience, location and role are concerned.

Interviews. A series of semi-structured interviews were conducted between December 2014 and February 2015 with self-selecting participants who completed the academic level survey. The purpose of the interviews was to explore experiences with some of the key themes and issues uncovered in the quantitative elements of the project. Interviews were conducted with 23 people from 15 different universities. Participants held a variety of roles (e.g. teacher, educational developer, student support officer, librarian, learning analytics project leader, tutor, and learning and teaching leader) and spanned different academic levels. Each interview was digitally recorded (audio) and then transcribed verbatim for coding and utilisation in the framework development phase. Data from the interviews was thematically analysed using manual coding by two researchers.

Triangulation Process. The two surveys were designed to include question blocks which were complementary to provide data on key themes emerging from the literature from different stakeholder perspectives. Data from these two sources were brought together to identify where further information was needed which informed the interview questions. Thematic analysis of the interviews was then brought together with the data from the two surveys to allow for manual comparison of the findings from each, identification of trends and key concerns. This was undertaken using a discursive approach by four team members. Through this process the headline findings were identified with data collated under each. Throughout this iterative process, themes, concerns, tensions and areas of focus were recorded to feed into the development of the framework and the associated question sets.

Phase 2

The framework and the associated question set were developed by the project team as a graphic representation of the key elements identified through phase 1 of the project. The first version was presented to the project reference group and their feedback incorporated. In the meantime, the project partners utilised the framework to develop case studies and tested it as a tool to inform their own institutional thinking.

The next step was to present the framework at a National Forum for feedback. This took place in April 2015, where 148 colleagues from 43 different institutions came together to discuss the findings, explore the framework and share examples from their own experiences with learning analytics. The project had a budget that allowed each Australian institution to send one fully supported (transport and accommodation) representative to the forum. Typically, this person was someone involved in implementing learning analytics, or who would be involved if learning analytics were to be implemented at their institution (e.g. DVC Academic, Director of Learning and Teaching, and Business Intelligence leaders). The rest of the participants made their own way to the forum, though registration was free. This resulted in a further group of stakeholders that included teaching staff, educational designers and developers and researchers, among others.

The morning session of the forum focused firstly on broad presentation of the project findings followed by a specific introduction to the framework (for details of the program and presentations see West et al, 2015a). The afternoon session split participants into two smaller structured workshop groups to explore and test the clarity and usefulness of the framework in the context of participants' own institutions and roles. During these workshops participants worked in small groups and members of the project team recorded key issues raised. Following the forum, the project team reconvened and amended the framework with further advice being provided by the project reference group.

Results

The results that follow are presented in two parts. The first of these is a series of 'headline' findings that spanned the three Phase 1 data sets (institutional level survey, academic survey and interviews). Given space limitations it is impossible to present all of the data tables and graphs but in this instance the headline findings are intended to lead into the presentation of the framework, which constitutes the second part of the results section. Readers interested in further exploring the data sets and findings can find this information on the project website (West et al, 2015a)

Headline Findings

Headline finding 1: The sector was at an early stage of development, implementation and understanding around learning analytics. Across the institution level survey, academic level survey and interviews, the data indicated that whilst pockets of learning analytics use existed in a number of institutions, there were many academics with very little awareness or participation in analytics discussions and projects. As an example, Figure 1 presents data from the academic level survey that illustrates this point. It shows results when participants were asked how often they discussed learning analytics with colleagues in different groups, with the darker bars on the right of each series showing less frequent discussion.

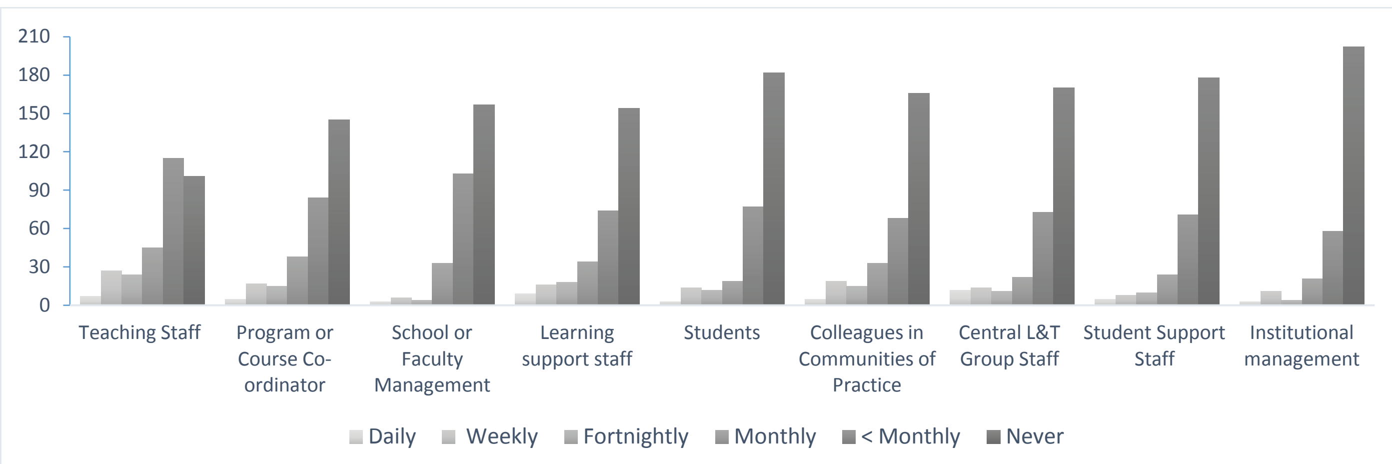


Figure 1: Frequency of discussion that participants reported with selected colleagues

The complexity of learning analytics may mean that development will take time and this may limit opportunities for academics to be involved and develop confidence with learning analytics in the short term. As one interviewee involved in learning analytics implementation stated: "I guess I am looking at where we might end up focusing on the spectrum of learning analytics and I don't really have an answer for that just yet". This is one quote that was indicative of many similar responses.

Headline finding 2: There was a degree of cautiousness reflecting an emerging recognition of the amount of thinking that needs to go into how learning analytics can be applied. Following on from the first headline finding, it became clear that perceptions differed on what learning analytics implementation might look like, as the following interview quote suggests:

> At our university, we're not too far advanced. We don't have that kind of system integration yet, that allows us to do that. Before we go headlong into assuming that it is all about system integration and spending huge amounts of money on trying to make systems talk to each other

> that don't want to, that if we drive the conversation on the basis of understanding what are the key points of data, it might be that those systems don't actually necessarily need to talk to each other, but we might be able to aggregate that information at another level.

This is where variation in leadership and strategic thinking can impact on what gets implemented and when. In some institutions there was a clear expression by those setting the direction that getting data into integrated systems, such as a data warehouses, would create the flexibility to address questions as they arise - one did not necessarily need to know what those questions were at the time of building the data warehouse. In contrast, other participants' views were more indicative of the quote above, which reflects a problem by problem approach in which implementation is driven by clearly identified needs and problems.

Headline finding 3: There was great variability across the sector – in relation to preparedness, how institutions are thinking about and implementing learning analytics due to context. The first two headline findings illustrate the complexity and broad scope of learning analytics and how this complexity posed challenges for participants in the study. The third headline finding follows logically but speaks to a different kind of complexity, which is the significant variation that can occur across institutional contexts and environments. This can take many forms, for example, variation in organisational culture as articulated by one interviewee:

> I think the real learning for me so far has been the need for the cultural environment at the institution and the associated drivers and leadership. The data has to have a certain environment or a certain context before it has any meaning whatsoever. That way you can clearly mark what that is and you can find a platform for it to have a sort of life within the environment.

Another commonly expressed example of variation in institutional context related to the student cohort. One participant described the student cohort at their university as "fairly competitive", another spoke of a large population of "low socio-economic" students, whilst another again spoke of their university being very "research focused". Taken together, these examples point to the importance of institutional context in shaping how learning analytics might be implemented and the framework unpacks this further.

Headline finding 4: Participants identified tensions and questions around the extent to which learning analytics can drive actions and behaviours or take the functions of people. Perhaps not surprisingly, given the Figure 1 data that illustrated the limited amount of analytics focused discussion taking place involving faculty, there was some uncertainty about how students would experience learning analytics, particularly automated communication. This tension was captured by one participant, who observed**:**

> Even though you are using automated systems to communicate and provide feedback you still have to make sure that it is personalised and meaningful for students and that takes time and consideration…In the system you can use this code so you don't have to write names every time and you don't have to use hello everyone all the time. This gets more attention from students and encourages students to be engaged.

Another participant offered the view that there would still be a major role for direct person-to-person communication with students.

> Statistics can be helpful, but it can also be useful to talk to students directly and see what they actually think they need themselves. From all of this we can work on programs, strategies and supports that might help them stick around a bit longer.

Whilst these two quotes express a consciousness of student needs they also potentially allude to a broader fear that one's job or work role might be impacted on, or even superseded, by computers.

Headline finding 5: There were substantial gaps between academics' needs and aspirations and what is being provided, what is available, and what is being developed. One of the challenges the sector faces in widely leveraging the potential of learning analytics is how to distribute knowledge and awareness of what can be achieved through learning analytics. As the previous quotes illustrated there are both real and potential fears around the implications of learning analytics for academic staff, particularly where academics are not well informed about what learning analytics are, how their institution plans to use them, and how they are expected to use them. The data in Figure 2 reinforces this point. Taken from the academic level survey, it shows that a large proportion of participants rated their institution as 'poor' or 'very poor' at meeting their needs in relation to some basic learning analytics provisions. Further illustrating the uncertainty felt by participants is that a relatively high proportion were unsure how they would rate their institution – *n* for 'not sure' varied between 85 and 104 across the six variables.

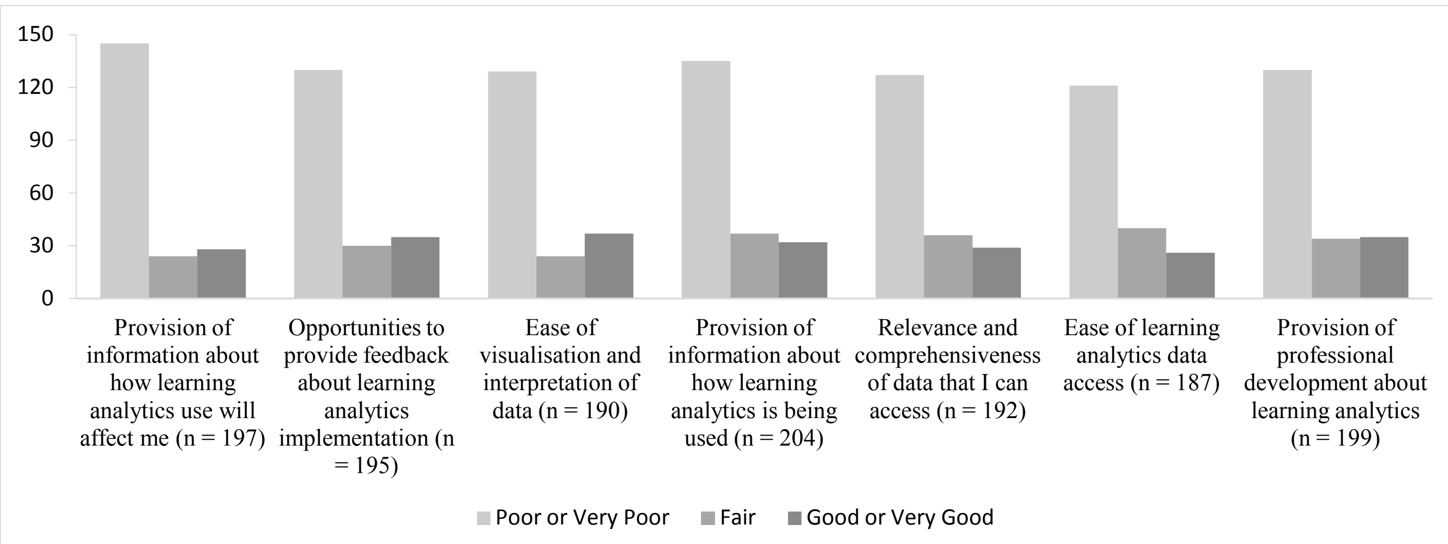


Figure 2: Participants rating of their institution at meeting their needs around selected learning analytic provisions (excludes 'not sure' responses to show trends visually)

The messages in this data are consistent with data provided by a number of interviewees as well as the institution level survey, which suggested many institutions were currently engaged in localised pilot projects and/or institution level systems development, which for various reasons may not be well publicised to the wider staff cohort in both cases. This leads to the final headline finding.

Headline finding 6: There was tension between business needs, wants and limitations (e.g. costs) versus academic staff needs and wants (e.g. academic freedom, innovation in learning and teaching). A recurring theme across the two surveys and interviews was that the roles of individual staff and work groups are a key factor in determining the types of learning analytics applications they might be interested in, and to some extent this is where the ambiguity around educational and learning analytics has impacts. There are two parts to this.

The first is that a number of participants expressed a desire for analytics to help inform their work, but there was also an expressed desire for academic freedom i.e. not being told how to teach, what forms grades had to be entered, or when to commence assessment tasks, for example. This can be problematic when particular processes need to be followed in order that data is cleanly and consistently recorded and thus usable. The second part relates to the beneficiaries of learning analytics. There are questions about the data that might be used for learning analytics, especially who owns it and how it can and should be used. When participants were asked about the ethical principles that should guide learning analytics, a number talked about the highest priority being student benefit. Others went as far as saying that the institution should not be using learning analytics for its own benefit.

Linking the Data and the Literature: A Summary of Lessons

These headline findings and a more granular thematic analysis of survey and interview responses (West et al, 2015) led the project team to consider how the various frameworks/models and benchmarking documents assist in implementation and address the issues described by participants. Key lessons from this reflection included:

- Despite a range of examples of learning analytics initiatives being available, people in the sector are struggling with institution level implementation of learning analytics, especially in terms of aligning policies, systems, people and stakeholder needs.
- Institutional context will change the shape of development and pathways for implementation
- A reflective, non-linear model is essential
- The domains outlined in the literature are essential but not sufficient, particularly in trying to address student success
- The models/frameworks are more broadly about technology enabled learning and/or learning analytics but are not focussed on the issue of student success
- Currently, there is a lack of clear alignment between the work of researchers, institutional learning analytics leaders and academic staff, particularly in terms of working together on problems of agreed strategic importance. The data in Figure 1 showed that academics in this sample were engaging in a relatively limited amount of dialogue with other colleagues, making strengthening dialogue a logical starting point.

Introducing the *Let's Talk Learning Analytics Framework*

With these assumptions in mind, the *Let's Talk Learning Analytics Framework* was developed (and refined as per the methodology above). It consists of two primary documents:

1. A one-page high level summary of the framework (Figure 3), which consists of six key domains. The six domains, shown in green, are the areas an institution needs to consider to make progress on learning analytics for retention. They are interconnected and shaped by the primary domain, the institutional context. Continual reflection and adjustment are critical to the process and will require input from across the institution.
2. A set of discussion questions (West et al, 2015c) to support the use of the framework, organised around the key domains. The purpose of the discussion questions is to provide a series of prompts to guide a contextualised discussion and is based on the premise that dialogue across an institution is critical.

The set of questions which accompany the framework are essential but, for this paper, will be split and presented under each of the domains to help illustrate the relevance of those domains. In practice it would be expected that learning analytics leaders would use the questions and framework documents side-by-side in a meeting, workshop, or any other environment centered on dialogue, or a combination of these in line with the yellow section in Figure 3. The questions are not statements related to stages or levels but rather, prompts, to engage in discussion and reflection to ensure that key elements are articulated, clarified or marked for development. The framework is also seen as dynamic and non-linear with continual reflection and adjustment, based on the idea of transitioning to learning analytics use. The term ‘transitional’ is used to reflect that learning analytics is an emerging field and many institutions will be transitioning to an environment that, for the first time, broadly incorporates learning analytics (e.g. into institutional policies and plans), as well as transitioning to an environment where learning analytics is specifically applied to issues like student retention. As such there is no implied point for starting the process beyond the institutional context. There is an expectation that institutions will engage in a process of praxis (reflection, action, reflection).

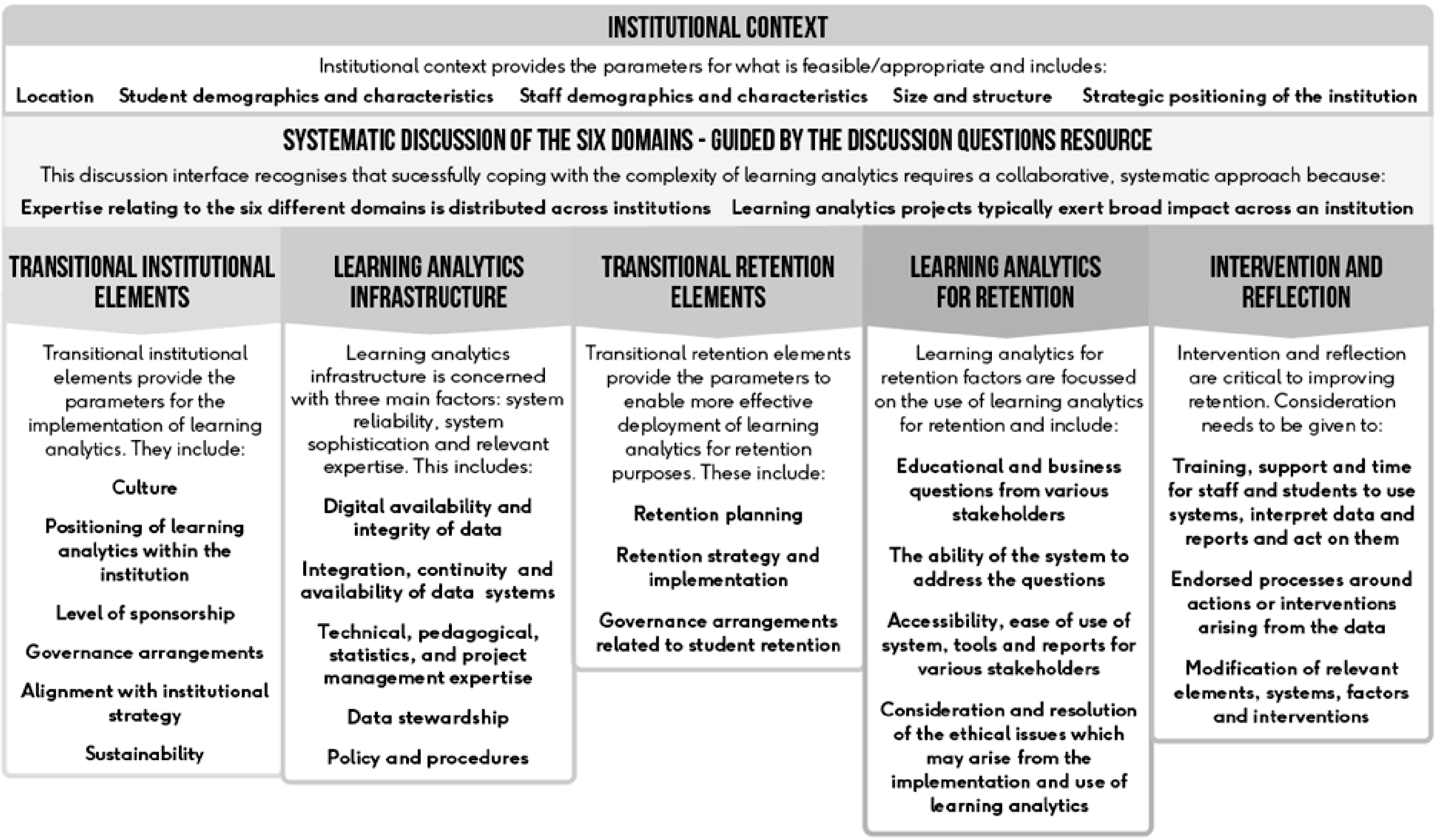


Figure 3: Let’s Talk Learning Analytics Framework

Domain 1: Institutional Context. The central message from the data was that institutional context is **the** guiding feature of analytics implementation. It sets the scene for all of the decisions that will follow and as such is seen as foundational and therefore positioned in the framework across the top to highlight the fact that it will pervade all other domains. The questions for consideration in this domain include:

- What are the implications of the jurisdictional location of your institution (e.g. privacy laws)?
- What is the size and structure of your institution?
- What is the strategic positioning of your institution (e.g. partnerships, international, and areas of focus)?
- What are the student demographics and characteristics?
- What are the staff demographics and characteristics?

Domain 2: Transitional Institutional Elements. Transitional elements in this framework are used as a term to both narrow the focus and to suggest that the elements that sit in these domains are a lens by which to view other domains and elements. They will inevitably shape the development of learning analytics development in an institution. For example, in institutional transitions, the executive commitment and position is critical to progress as it will impact on investment, practices and culture (see Arnold et al, 2014 and Norris & Baer, 2012). Key questions for discussion in this domain include:

- How does the culture at your institution both promote and challenge the advancement of learning analytics?
- What is the strategic positioning of learning analytics within your institution?
- To what degree is there executive sponsorship (academic and IT)?
- Do you have a specific learning analytics plan and if so, how is it aligned with your institutional budget and academic plan?
- To what extent is there congruence between the vision, strategy and values of the institution that relate to learning analytics?
- How clearly is the role of learning analytics in your institution identified and communicated?
- How sustainable is the approach to learning analytics at your institution?

Domain 3: Learning Analytics Infrastructure. The Learning Analytics Infrastructure domain is a representation of the importance of data systems and infrastructure to the use of learning analytics. There is no doubt that institutions have been collecting and storing significant amounts of data for years (Educause, 2010; Heathcote & Dawson, 2005). However, these have not necessarily been accessible in the right format or able to be used effectively. Project participants cited numerous examples of issues around data and systems where the interface is not user-friendly, the data is incomplete, data integration is inadequate, or the time taken to obtain a report or visualisation is too long.

This domain includes three key questions which are then subsequently broken down into more specific questions under the headings of digital availability of data, data integrity, systems continuity, availability and reliability, integration of data systems, expertise, stewardship and data policy and procedures. For our purposes only the three key questions are presented here (the fuller set area available on the project website (West, 2015c).

- Are your systems reliable?
- Are they technically sophisticated enough?
- Is there enough expertise in the institution?

Domain 4: Transitional Retention Elements. As this framework is specifically focussed on the use of learning analytics for student retention, this domain draws attention to that lens and transitions us from thinking about all the things learning analytics might be used for, to how it can be used to support student success. Key questions in this transition domain include:

- Do you have an institutional retention plan or strategy?
- What are the governance arrangements that support the retention plan or strategy?
- How are learning analytics positioned within the retention plans and associated governance arrangements?

It is important to reiterate that this framework was developed for a project with a specific focus on student retention. This explains the focus of Domains 4 and 5. It is possible to adapt these domains to other key foci by changing the focus of Domains 4 and 5, whilst keeping the other domains as they are.

Domain 5: Learning Analytics for Retention. This domain highlights key questions that prompt thinking around the more specific area of learning analytics for retention (rather than business intelligence, action analytics etc.) and so should be considered in that light.

- What business and educational questions do your stakeholders have, and how are they prioritised?
- To what extent can your system address those questions?
- Where gaps are identified, what resources are available for follow up?
- What ethical issues have been identified and what is the resolution?

Domain 6: Intervention and Reflection. Data gathered throughout the project demonstrated that there is a wide range of learning analytics activities taking place throughout the sector. However, many of these are small scale or pilot projects are not necessarily centrally led and take place in various locations with different intended outcomes. Yet, projects with no action will have little impact on student success and they thus need to be connected to some sort of intervention (whether the data informs an action (utilising people to call students) or be part of the action (e.g. automated messages to students or connections to personalised learning). Intervention at this stage also includes things like training programs for staff and students, which was clearly identified as a gap in the data.

It was also clear that while reflection is necessary in an ongoing manner, it needed to be more explicit in this area to ensure that lessons from learning analytics projects are fed back in to other work. A clear example of this might be that a learning analytics project demonstrates that a particular strategy in the retention plan is not effective. This should prompt investigation and subsequent revision of the retention plan. Key questions in this domain include:

- How are potential interventions developed, managed and enabled?
- What training is provided, and what training is required for staff and students?
- What time and support is provided, and what time and support are required, for staff and students?
- How are the interventions evaluated and improved?
- How is the whole system evaluated and improved?

Discussion

While this framework builds on prior work (Arnold, Lonn, et al., 2014; Bischsel, 2012; Norris & Baer, 2012), it has some key differences. These include an emphasis on dialogue, a focus on the use of learning analytics for student retention and success, and placing an explicit focus on both the foundation of context and the action/intervention required. It attempts to draw attention to the inter-related elements that will take learning analytics in different, but equally useful directions.

Clearly the level of maturity and capacity that institutions have when it comes to learning analytics varies and though there has been much research interest in learning analytics, this is not necessarily translating into coherent, widely understood plans and strategies within institutions. The flow-on effect of this is that academic staff may be anxious about falling behind with a new set of tools (i.e. learning analytics) that they keep hearing about. Additionally, they may not be aware that their knowledge and skills could be harnessed in relation to institutional learning analytics work. In this way, the framework is seen as a key tool for creating constructive dialogue around the potentially daunting changes that learning analytics bring.

The Role of Dialogue

The emphasis on dialogue hinges on several factors that were very evident in both the literature and the data. First, that learning analytics is complex and in a nascent stage of development, which means that we are still exploring the potential that it affords. In order to explore and leverage the potential, we need to have the involvement of people from multiple disciplines and backgrounds (Balacheff & Lund, 2013; Siemens, Dawson & Lynch, 2013) (but particularly IT and pedagogy), and points of view (such as teaching academics, head of schools, senior executive and students) at every stage. This includes planning, development, deployment, and feedback. Yet the data suggested that such knowledge and input was largely occurring in fairly narrow silos or select groups, and teaching academics in particular were mostly unaware of the initiatives taking place in their own institutions. As noted in Figure 1, most academics in the survey very rarely had discussions with anyone around learning analytics. This is not particularly surprising given that many initiatives were small pilot projects of limited scope. Additionally, the complexity of learning analytics initiatives from a technical view has placed it largely in the realm of IT departments or learning analytics specialists. Breaking down the potential silos is critical and dialogue is one way to do this.

Furthermore, many people (even at higher levels) in the sector have had limited practical exposure to learning analytics and therefore do not always recognise tangible implications that impact on implementation. Finally, learning analytics implementation necessitates collaboration between people with vastly different roles and expertise. For example, the use of even basic reporting for functions from software packages which integrate the LMS and SIS, need a level of standardisation in practice to be effective. Items such as the grade centre need to be set up in an appropriate way in order to display some of the reports. Conversations about these factors and which reports would be most useful from the perspective of the teaching academic, could help in a range of ways including to prioritise work in analytics implementation; aid in a level of standardised practice by teaching academics; and, link to support mechanisms for staff and students.

Bichsel (Bischsel, 2012) recommends that institutions use a maturity index (specifically the Educause maturity index) to gauge analytics progression, set goals and measure progress. While this type of benchmarking is useful, the idea of plotting progression on a scale is generally of interest to particular stakeholders. Additionally, it is not designed to be an implementation tool or guide but it is about measuring progress instead. In contrast the framework presented here is not about benchmarking but about providing a set of discussion questions to be actively used to move implementation (or at least dialogue on analytics) forward. This is very much in line with the suggestions of Bichsel (2012) who recommends attention to planning, process and partnership across various stakeholders.

Analytics for Student Retention and Success

As noted at the outset of this paper, the area of student retention and success is a high priority in the sector. This focus has flowed through to the use of learning analytics being harnessed to assist in this area (for example see (Arnold & Pistilli, 2012; Sclater, 2014; Yeado et al., 2014). Yet frameworks and benchmarking tools on analytics are about analytics more broadly, which, while useful, does not prompt stakeholders to think about the complex area of student retention nor to frame the work around the broader context of student success. While perhaps there is an assumption that that we would be using learning analytics to improve the student journey, there is value in making it more explicit.

While the authors acknowledge and support the idea that learning analytics can be used for a wide range of purposes, central to this is the idea of student retention and success. This project was funded to consider learning analytics in light of retention but it was very apparent that thinking in the sector has moved to the broader construct of student success with retention being a subset of that. In many ways a clear focus on the student journey may help to move the learning analytics journey forward. As Siemens (2012: 4) observes: “A transition in analytics from a technical orientation to one that emphasizes sense-making, decision-making, and action is required to increase interest among educators and administrators. This transition in no way minimizes the technical roots of analytics; instead, it recasts the discussion to target problems that are relevant to practitioners instead of researchers”. It is perhaps the student journey which may be able to provide a point of reference for engagement of a more diverse set of stakeholders.

The work in the area of student retention can also provide some guidance in terms of the types of initiatives which may be useful, and this will flow through to the types of data that would be collected, the types of reports required, who should have access to those reports and the subsequent actions that might take place. By way of example, Transition Pedagogy (Kift, Nelson & Clarke, 2010; Kift, 2009; Kift & Nelson, 2005) focuses on the concepts of transition, diversity, design, engagement, assessment and evaluation as being critical to first year retention. If an institution was using such a framework it would be much clearer on how to focus learning analytics development. Additionally, some of the concepts within Transition Pedagogy are notoriously difficult to measure and learning analytics may offer some opportunities to revisit such work to both test and expand the model.

Context

Context is critical. This basic statement was a central theme of the data in terms of the initiatives occurring, what was possible, the resources and infrastructure available and the priorities set to name just a few. In this sense, case studies and examples of what has worked in other places are useful but not sufficient. A good example of context within a domain is the strategic positioning of learning analytics within an executive portfolio. In some institutions this sits within an academic portfolio and in others in the IT portfolio. Such position in one or the other is both a reflection of how the institution views learning analytics and its purpose, as well as a driving force for future direction. Starting with the context will ensure that the initiatives that take place are appropriate to that environment or at least considered in this light.

Leveraging analytics is a dynamic and interactive process. There are a range of implementation decisions which need to be made along the way. There is the potential for these to be driven by a narrow focus rather than an institutional one, potentially leading to a mismatch between the decision and the direction of the institution, or, if present, the retention plan. For example, the decision to turn on a student facing dashboard is deceptively simple but may require deeper consideration in some contexts. Consideration may need to be given to the support that is provided around this, the nature of the report or dashboard, and the type of follow-up that takes place. Additionally, such dashboards should be built in relation to key pedagogic or retention questions. It is essential that each one of these decisions is in line with the overall plan and approach, and seen in context. The point is that there is no simple right or wrong but rather decisions that are built in line with institutional elements.

Finally, the decision to pursue learning analytics is costly in terms of time, expertise and money. If the decision and subsequent implementation is done in a systematic way and includes various stakeholders and expertise in development, it is more likely to have an impact. A number of participants emphasised that inadequate communication, where learning analytics implementation is concerned, is liable to cost time and money. Therefore, from an efficiency point of view the use of a dialogical framework can prompt the communication in a focussed way and avoid unnecessary delays and align processes for outcomes.

Limitations

Prior to the concluding remarks some potential limitations are acknowledged:

- Learning analytics is rapidly developing area so it is likely that ongoing data collection is required to gather updated perceptions, experiences and aspiration from academics and institutions as they get more experience with learning analytics. To address this limitation, the project team plans to conduct follow up data collection in the coming period.
- The data used to develop the framework was from Australia and New Zealand, so readers will need to consider how this might impact on the applicability of the framework in their location, or other locations. However, because it is a discussion framework the researchers are interested in hearing feedback from peers in other locations and this will potentially assist in further refinement.
- The process of developing and refining the framework via data and various forums was highly iterative and describing the nuances of that method for the purposes of replicability is a challenge. Questions can be directed to the researchers.
- In terms of this specific paper it must be acknowledged that space restrictions also limited the amount of raw data that could be presented.

Conclusion

This paper has presented findings from a mixed-method study of Higher Education institutions and academics from Australia and New Zealand. Whilst this project produced a range of data on the perceptions, experiences and aspirations around learning analytics, one of the major aspirations of the project was to translate this data into a framework to promote institutional implementation bolstered by robust and systematic discussion. This has taken the form of the *Let's Talk Learning Analytics Framework*, and its presentation has been the primary focus of this paper. The framework can be thought of as comprising two complimentary tools. The first is the map of domains and factors relevant to institutional implementation of learning analytics and the second is a series of discussion questions that act as prompts within each of the domains.

Overall the framework is designed to foster a collegial approach to the implementation of learning analytics which is based on providing prompts for dialogue amongst key stakeholders. In this way it compliments existing frameworks and models which are more oriented toward measuring progress. Ultimately, it is hoped that this will assist in positioning learning analytics in a shared space which equally values pedagogical and technical expertise. Most importantly, it suggests that a focus on the student journey and the associated operational plans of an institution will be more likely to connect analytics to action for the benefits of students.

In the same way that it takes a community to raise a child, it takes an institutional community to truly make the most out of learning analytics.

Acknowledgements

This project was funded by the Australian Government Office for Learning and Teaching. The views expressed in this publication do not necessarily represent those of the Australian Government Office for Learning and Teaching.

The authors would like to acknowledge that this framework was developed by the entire project team which included the authors and Professor A Lizzio, Griffith University; Professor C Miles, Newcastle University; Dr J Bronnimann, Batchelor Institute of Indigenous Education, Mr D Toohey, Murdoch University and Mr B Searle, Charles Darwin University.

References

ACODE. (2014). *Benchmarks for Technology Enhanced Learning*. Canberra, ACT: ACODE.

Arnold, K., Lonn, S. & Pistilli, M. (2014). An exercise in institutional reflection: The Learning Analytics Readiness Instrument (LARI). Proceedings *5th International Conference on Learning Analytics and Knowledge Conference (LAK'15)*, Poughkeepsie, New York.

Arnold, K., Lynch, G., Huston, D., Wong, L., Jom, L. & Olsen, C. (2014). Building institutional capacities and competencies for systemic learning analytics initiatives. Proceedings *Learning Analytics and Knowledge Conference (LAK'15)*, Poughkeepsie, New York.

Arnold, K. & Pistilli, M. (2012). Course signals at Purdue: Using learning analytics to increase student success. Proceedings *2nd International Conference on Learning Analytics and Knowledge (LAK '12),* Vancouver, British Columbia.

Balacheff, N. & Lund, K. (2013). Multidisciplinarity vs. multivocality: The case of "Learning Analytics". *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge (LAK '13),* Leuven, Belgium. doi:10.1145/2460296.2460299

Behrendt, L., Larkin, S., Griew, R., & Kelly, P. (2012). *Review of higher education access and outcomes for Aboriginal and Torres Strait Islander People: Final report*. Canberra: Australian Government.

Bichsel, J. (2012). *Analytics in higher education: Benefits, barriers, progress and recommendations*. Louisville, CO: EDUCAUSE Center for Applied Research.

Bradley, D., Noonan, P., Nugent, H. & Scales, B. (2008). *Review of Australian higher education: Final report.* Canberra: Australian Government.

Buckingham Shum, S. & Ferguson, R. (2012). Social learning analytics. *Educational Technology & Society*, 15(3), 3-26.

Campbell, J., and Oblinger, D. (2007) *Academic analytics: White paper.* Educause. Retrieved 12 May, 2015 from http://net.educause.edu/ir/library/pdf/pub6101.pdf

Colvin, C., Rogers, T., Wade, A, Dawson, S., Gasevic, D., Buckingham Shum, S., Nelson, K., Alexander, S., Lockyer, L., Kennedy, G., Corrin, L., & Fisher, J. (in press). *Student retention and learning* *analytics: a snapshot of Australian practices and a framework for advancement*. Strawberry Hills, NSW: Office for Learning and Teaching.

Dekker, G., Pechenizkiy, M. and Vleeshouwers, J. (2009). Predicting students drop out: A case study. Proceedings *2nd International Conference on Educational Data Mining (EDM'09),* pp. 41-50.

Educause (2010). *7 Things you should know about analytics.* Retrieved May 12, 2015 from http://www.educause.edu/ir/library/pdf/ELI7059.pdf

Educause Center for Applied Research (ECAR) (2012). *Analytics Maturity Index for Higher Education*. Retrieved July 21, 2015 from http://www.educause.edu/ecar/research-publications/ecar-analytics-maturity-index-higher-education

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. International *Journal of Technology Enhanced Learning*, 4(5/6), 304-317.

Fiaidhi, J. (2014). The next step for learning analytics. *IT Pro*, September/October, 4-8.

Frankola, K. (2002). *Why online learners drop out.* Retrieved 14 June 2013 http://www.kfrankola.com/Documents/Why%20online%20learners%20drop%20out_Workforce.pdf

Gaševic, D., Dawson, S. & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends: Linking Research and Practice to Improve Learning*, 59(1), 64-71.

Heathcote, E. & Dawson, S. (2005). Data mining for evaluation, benchmarking and reflective practice in a LMS. Proceedings *e-Learn 2005: World conference on e-Learning in corporate, government, healthcare & higher education*, Vancouver, Canada.

Johnson, L., Adams Becker, S., Estrada, V. & Freeman, A. (2015). *NMC Horizon Report: 2015 Higher Education Edition*. Austin, TX: New Media Consortium.

Kift, S. (2009). *Articulating a transition pedagogy to scaffold and to enhance the first year student learning experience in Australian higher education*. Strawberry Hills, ACT: Australian Government (ALTC).

Kift, S. & Nelson, K. (2005). Beyond curriculum reform: Embedding the transition experience. In Brew, A. & Asmar, C. (Eds.), *Higher Education in a changing world: Research and Development in Higher Education* (pp. 225-235). Milperra, NSW: HERDSA.

Kift, S., Nelson, K. & Clarke, J. (2010). Transition Pedagogy: A third generation approach to FYE – A case study of policy and practice for the higher education sector. *The International Journal of the First Year in Higher Education*, 1(1), 1-20.

Koshy, P. (2014). *Student equity performance in Australian higher education: 2007 to 2012.* National Centre for Student Equity in Higher Education (NCSEHE), Perth: Curtin University.

Long, P. & Siemens, G. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, 46(5), 30-32.

Marshall, S. (2006) *eMM Version Two Process Guide*. Wellington: Victoria University.

Marshall, S. & Mitchell, G (2002). An E-Learning Maturity Model? Proceedings *19th Annual ascilite Conference*, Auckland, New Zealand.

Nelson, K., Clarke, J., Stoodley, I. & Creagh, T. (2014). *Establishing a framework for transforming student engagement, success and retention in higher education institutions*. Final Report 2014, Canberra: Australian Government Office for Learning & Teaching.

Norris, D., & Baer, L. (2012). A Toolkit for Building Organizational Capacity for Analytics (pp. 1-180). Washington, DC: Strategic Initiatives Inc.

Rivera, J., & Rice, M. (2002). A comparison of student outcomes and satisfaction between traditional and web based course offerings. *Online Journal of Distance Learning Administration*, 5(3).

Sclater, N. (2014). *Learning analytics: The current state of play in UK higher and further education*. JISC, United Kingdom.

Siemens G. (2012). Learning analytics: Envisioning a research discipline and a domain of practice, Proceedings *2nd International Conference on Learning Analytics and Knowledge (LAK '12),* Vancouver, British Columbia.

Siemens, G., Dawson, S. & Lynch, G. (2013). *Improving the productivity of the higher education sector: policy and strategy for systems-level deployment of learning analytics*. Canberra, ACT: SoLAR/Australian Government Office for Learning and Teaching.

Universities Australia (2013). *A smarter Australia: An agenda for Australian higher education 2013-2016.* Canberra, ACT: Universities Australia.

van Barneveld, A., Arnold, K., & Campbell, J. (2012*). Analytics in higher education: Establishing a common language.* Educause Learning Initiative. Retrieved May 10, 2015 from http://net.educause.edu/ir/library/pdf/eli3026.pdf

West, D., Huijser, H., Heath, D., Lizzio, A., Miles, C., Toohey, D., Brönnimann, J. & Searle, B. (in press). *Learning analytics: Assisting universities with student retention*. Strawberry Hills, NSW: Australian Government Office for Learning and Teaching.

West, D., Huijser, H., Heath, D., Lizzio, A., Miles, C., Toohey, D., Brönnimann, J. & Searle, B. (2015a). *Forum resources*. Retrieved September 25, 2015 from http://www.letstalklearninganalytics.edu.au/national-forum/

West, D., Huijser, H., Heath, D., Lizzio, A., Miles, C., Toohey, D., Brönnimann, J. & Searle, B. (2015b). *Discussion Questions*. Retrieved September 25, 2015 from http://www.letstalklearninganalytics.edu.au/wp-content/uploads/2015/06/Discussion-Questions-Final-26-May-2015.pdf

West, D., Huijser, H., Heath, D., Lizzio, A., Miles, C., Toohey, D., Brönnimann, J. & Searle, B. (2015c). *Let's Talk Learning Analytics Framework*. Retrieved September 25, 2015 from http://www.letstalklearninganalytics.edu.au/wp-content/uploads/2015/06/Lets-Talk-Learning-Analytics-Framework-Final-26-May-2015.pdf

Willcoxson, L., Manning, M., Hibbins, R., Joy, S., Thomas, J., Girardi, A., Leask, B., Sidoryn, T., Cotter, J., Kavanagh, M., Troedson, D., & Lynch, B. (2011). *The whole of university experience: Retention, attrition, learning and personal support interventions during undergraduate business studies*. Surrey Hills, NSW: Australian Learning and Teaching Council. Retrieved May 10, 2015 from http://eprints.usq.edu.au/20138/

Yeado, J., Haycock, K., Johnstone, R., & Chaplot, P. (2014). *Education trust higher education practice guide: Learning from high-performing and fast-gaining institutions*. The Education Trust, Washington, DC.

Assessing the Reliability of Merging Chickering & Gamson's Seven Principles for Good Practice with Merrill's Different Levels of Instructional Strategy (DLISt7)

Dr. Syaril Izwann Jabar
University Selangor, Malaysia

Prof. Peter R. Albion
University of Southern Queensland, Australia

Abstract

Based on Chickering and Gamson's (1987) Seven Principles for Good Practice, this research project attempted to revitalize the principles by merging them with Merrill's (2006) Different Levels of Instructional Strategy. The aim was to develop, validate, and standardize a measurement instrument (DLISt7) using a pretest-posttest Internet quasi-experiment. It was proposed that the instrument could then be used as a rubric either for facilitating the implementation of DLISt7, or as a set of unobtrusive diagnostic indicators for assessing the quality of learning experienced by students in blended and online courses. The study was conducted across five faculties at a regional Australian multi-campus university. The intent was to contribute to knowledge building by leveraging the data that had been collected, analyzed, and reported to generate awareness about the likelihood of scaffolding and scaling, varying levels of instructional strategies for communicating expectations, and relaying information. The idea was to produce a tool that would create more opportunities for more of the principles to be put to good use as an effectiveness multiplier. The findings from the analysis conducted using exploratory and confirmatory factor analysis verified the validity of DLISt7 and demonstrated excellent internal reliability values.

Introduction

In the same way discoveries and inventions fuelled the Industrial Revolution mechanizing human society, technology has often been credited as the catalyst that modernized the way learning can be organized and teaching resources utilized (Lever-Duffy, McDonald & Mizell, 2003; Shneiderman, 1998). However, the debate continues about how successful academics have been at leveraging what computers, mobile devices and the Internet have to offer in terms of modernizing the delivery of education to learners (Lever-Duffy & McDonald, 2008). Despite the enthusiastic adoption of blended and online learning in higher education, research into web-based methodologies have revealed contradictory results that have not yielded clear pedagogical ground rules (Garrison, Cleveland-Innes & Tak Shing Fung, 2010).

Thus, there is little value in knowing what contemporary technology has to offer when educators are hesitant about when, where and how to best use technology to support the process of learning and teaching (Syaril Izwann Jabar, 2012a). Chickering and Ehrmann (1996) noted that simply having access does not guarantee being able to effectively leverage and innovatively utilize what contemporary technology has to offer, henceforth referred to as instructional technology (IT) so as to include the use of instructional media for the process of instruction (Morrison, Ross & Kemp, 2001; Reiser, 2012).

The challenge of utilizing IT to support a wide range of pedagogical approaches borrowed from the conventional classroom is already a difficult undertaking (Albion & Redmond, 2006), and becomes even more challenging when having to make it all work in an effective, efficient and engaging manner (Merrill, 2006). Even though technology often takes centre stage, it is actually the connections made with humans that drive online learning systems. Thus, the fundamental issue that arises for online education is "how to marry the power of networked connectivity with established pedagogical principles to produce better learning outcomes" (Kehrwald, Reushle, Redmond, Cleary, Albion & Maroulis, 2005, p. 1).

Although the technologies used to facilitate learning have evolved from classroom blackboards to www.blackboard.com, the characteristics of teaching and learning have not really changed; good teaching is still good teaching (Albion & Redmond, 2006) but "what has changed is how education providers and teachers facilitate learning" (Kehrwald et al., 2005, p. 1). For example, one widely recognised formulation of such principles is the Seven Principles for Good Practice in Undergraduate Education (Seven Principles) that was first proposed for face-to-face education (Chickering & Gamson, 1987) and subsequently adapted for a variety of contexts and purposes (Chickering & Gamson, 1999). The principles affirm that good practice;

1. *Encourages contact between students and teaching staff*
2. *Develops reciprocity and cooperation among students*
3. *Encourages active learning*
4. *Gives prompt feedback*
5. *Emphasizes time on task*
6. *Communicates high expectations*
7. *Respects diverse talents and ways of learning*

Teachers working in online environments can transfer what they know about teaching from other environments but need to consider how to adapt their knowledge while managing technology to balance flexibility and structure, put the learner at the centre of a supportive process, and create opportunities for engagement and interaction that stimulate learning (Kehrwald et al., 2005).

The complex nature of online instructional design suggests the need for practical ground rules designed to enable the effective integration of instructional technology with good pedagogy regardless of form, delivery system, or instructional architecture (Merrill, 2008; Merrill, 2009). Teaching staff with traditional, minimal or no online teaching experience should not have to rely on the process of trial and error to learn about how to teach effectively while at the same time scrimmage with the idiosyncrasies of the modern online environment (Haughton & Romero, 2009).

Clearly, there is a need for some form of guidance that educators can adapt to their personal style, course content, student population, and available technology (Shneiderman, 1998). One source of such guidance is the Different Levels of Instructional Strategy (DLIS) proposed by Merrill (2006). He describes five levels of application based on his First Principles of Instruction, namely, Demonstration, Application, Task-centred, Activation, and Integration. Progress along this continuum can be charted using four levels (0-3), Level 0 is Information only, Level 1 is Information plus Demonstration, Level 2 is Information plus Demonstration plus Application, and Level 3 is Task-centred integration of all three preceding levels (Merrill, 2006, 2009).

Of particular concern is the possibility of a gap in the synergy of events between cognitive presences, social presences, teaching presences, and strategies or tactics for online learning and teaching (Kehrwald et al., 2005). For example, a recommendation was made on the back of findings from a number of research projects about the widely accepted Community of Inquiry (CoI) framework that learning presence should be assimilated as a new conceptual element (Shea, Hayes, Smith, Vickers, Bidjerano, Pickett, Gozza-Cohen, Wilde & Jian, 2012).

Within the context of this research a similar recommendation was made with regards to issues associated with innovation complexity, obscurity of results, and lingering doubt about the effectiveness, efficiency, and engagement of online education (Syaril Izwann Jabar, 2007). Consequently, in order to maximise the benefits of online learning, learners need to be encouraged to accommodate strategies such as forethought, planning, mentoring, learning performance and reflection that good students are supposed to utilize while self-regulating learning presence (Shea, Hayes, Smith, Vickers, Bidjerano, Gozza-Cohen, Jian, Pickett, Wilde & Tseng, 2013).

Thus, it is proposed that in order for the science of learning and the art of teaching (Skinner, 1968) to be more effective in blended or online environments, the eclectic selection of appropriate pedagogy should consider the systematic and conscientious use of contextual engagement. This is because, as the world's population continues to grow "a far greater part of them want an education" for which "the demand cannot be met simply by building more schools and training more teachers" (Skinner, 1968, p. 29).

Instead, "education must become more efficient. To this end curricula must be revised and simplified, and textbooks and classroom techniques improved" upon using the latest available technology (Skinner, 1968, p. 29). Teachers for instance, can help students "in every way short of teaching them" by communicating new approaches about the acquisition of knowledge, not to mention the relaying of information about the "methods and techniques of thinking, taken from logic, statistics, scientific methods, psychology, and mathematics. That's all the 'college education' they need. They [can] get the rest by themselves in our libraries and laboratories" (Skinner, 1962, p. 121).

Literature Review

Integrating Instructional Technology with Good Pedagogy

As it stands, pedagogy is the action of teaching, or what teachers do when implementing their craft to assist students' learning (Lever-Duffy, McDonald & Mizell, 2005). In an age when information is expanding and becoming more readily accessible, the most critical outcome of education is learning how to learn. The application of sound educational principles and methods that informs teachers' practice can perhaps lead to improvements in the sequencing of educational experiences for designing, developing, implementing, and evaluating a pedagogically sound learning environment that is conducive for advancing knowledge building (Lever-Duffy & McDonald, 2008).

An understanding of learning and teaching has developed alongside technology and generations of educators have recognised that "as a mere reinforcing mechanism" the teacher might one day become antiquated, but "if the teacher is to take advantage of recent advances in the study of learning, she must have the help of mechanical devices" (Skinner, 1968, p. 22). Such an idea was probably thought of at that point in time as science fiction because nobody yet understood how mechanical and electrical devices would evolve into modern one-to-one, one-to-many or many-to-many robust supporting communication technologies that could be skilfully used to progress through a series of successive shaping approximations that shortened the window of opportunity between reinforcement contingencies (Anderson & Dron, 2012; Case & Bereiter, 1984; Skinner, 1954). That was true for the blackboard in its time and continues to be true for the evolution from Skinnerian behaviourism to cognitivism and for successive generations of technology enhanced distance education because of the awareness that "the human organism is, if anything, more sensitive to precise contingencies" (Skinner, 1954, p. 94).

The Digitization of Education

Scientific advancement is not a smooth evolutionary process, but is more like a sequence of peaceful intervals interrupted by intellectually violent upheavals in which an individual's attributional interpretation of the world is improved upon or replaced by another (Kuhn, 1970). Educational change exhibits similar characteristics and perhaps the time has come to re-evaluate the current paradigm of online pedagogy and set new directions for continued growth, modernization, and eventual maturity.

In light of the fact that "transitioning from teaching in the traditional classroom to the online environment is not a simple task for most faculty," there exists the need for a set of ground rules that can be used to improve the whole experience and enable the continuation of good teaching practices integrated with instructional technology (Grant & Thornton, 2007, p. 352). Anderson and Dron (2011) articulated the idea best when they said that quality online learning experiences make the most of cognitive-behaviourist, social constructivist and connectivist pedagogies across three generations to encapsulate what distance education has evolved and matured into.

Theoretical Rationale

Learner engagement can be defined as the quality of effort, in terms of time and energy invested in purposeful interactive involvement in educational activities and conditions that are likely to contribute directly to the construction of understanding (Coates, 2006; Kuh & Hu, 2001). When achieved alongside effectiveness and efficiency, engagement can lead to an increase in learner interaction, interest, and subsequently satisfaction (Merrill, 2008; Merrill 2009). According to Bangert (2008a), "a recommended instructional practice for higher education faculty is to engage in valued and meaningful inquiry-based, learning activities" that are "designed to create knowledge structures that can be retrieved when necessary to solve problems encountered in real-world contexts" (p. 36).

An aim of online education has been to actualize the potential afforded by communication and Internet technologies via design options that would enable participants to maintain engagement in a community of learners using asynchronous interaction (Garrison & Cleveland-Innes, 2005). Another goal has been to "structure the educational experience to achieve defined learning outcomes" using interaction that is structured, systematic, critical, and reflective (Garrison & Cleveland-Innes, 2005, p. 134).

Analyzed closely, these are also the achievable learning and teaching goals that are the focus of the Seven Principles (Chickering & Gamson, 1987) and DLIS (Merrill, 2006, 2009). The outcome of an education programme is often simply based on the premise of encouraging the active construction and demonstration of understanding, or as stated by Scardamalia and Bereiter (2006) "All understandings are inventions; inventions are emergents" (p. 15). Hence, as described in this paper the Different Levels of Instructional Strategies (DLISt7) for Online Learning was designed to function intelligibly as, "a set of workable principles that could guide pedagogy in a variety of contexts" (Scardamalia & Bereiter, 2006, p. 24).

Research Design

Statement of the Problem

This study was designed to investigate whether the Seven Principles could be revitalised by merging them with DLIS to form DLISt7. The approach was to develop a measurement instrument that could be used either as a rubric for facilitating the extrinsic implementation of DLISt7, or as a set of unobtrusive diagnostic “process indicators” (Kuh, Pace & Vesper, 1997, p. 436) for intrinsically assessing the quality of learning experienced by students in blended and online courses.

The intent was to contribute by leveraging the data that had been collected, analyzed, and reported to generate awareness about the likelihood of scaffolding and scaling varying levels of instructional strategies to make informed improvements in the instructional design of future online courses. The idea was to produce a tool that would create more opportunities for more of the principles to be put to good use as an effectiveness multiplier in support of efficient and engaging online learning. The critical insight for educational administrators, teaching staff, and instructional designers is the importance of utilizing learning analytics to make informed decisions about the appropriate balance between the eclectic utilization of asynchronous or synchronous communication technology and other available online resources.

In other words, when DLIS is used as a rubric either for teaching or treatment purposes to prompt and stimulate conditional response from students - which explains the *t* in DLISt7 - favourable online learning experiences that are consistent with the Seven Principles would be realistically achievable in ways that are familiar and unobscure (Syaril Izwann Jabar, 2012c).

Focus of the Research

An instructional principle, as defined in the context of this study, is “a relationship that is always true under appropriate conditions regardless of the methods or models which implement this principle” and whose underlying function is “to promote more effective, efficient, or engaging learning” (Merrill, 2009, p. 43). In their original form, the Seven Principles were designed to be robust so as to always be true under appropriate conditions with each principle having the capacity to “stand alone on its own, but when all are present their effects multiply” (Chickering & Gamson, 1987, p. 2).

Upon being updated, the term “instructional strategy” was integrated so as to accentuate the utility of the Seven Principles in promoting effective, efficient, and engaging learning in conjunction with “new communication and information technologies that had become major resources for teaching and learning in higher education” (Chickering & Ehrmann, 1996, p. 1). Despite their apparent simplicity and practicality, there has been a tendency for the Seven Principles to not be fully utilized (Bangert, 2004; Bangert 2008b; Batts, 2008; Chickering & Gamson, 1999; Cobbett, 2007, & Wuensch, Shahnaz, Ozan, Kishore & Tabrizi, 2009).

A review of the above literature suggests a penchant for the Seven Principles to be implemented and subsequently assessed in their standalone form instead of as a whole. Perhaps the Seven Principles could be resuscitated by being analysed from a different perspective. To echo the words of Merrill, “we need to back up and find out if there's a set of principles we can agree to and then build on these principles. Let's build on what's there instead of starting over and reinventing the wheel every single time” (as cited in Spector, Ohradza, Van Schaack & Wiley, 2005, p. 318). Thus, this study was devised to approach the Seven Principles from a different perspective, that is, by linking them with DLIS (Chickering & Gamson, 1987; Merrill, 2006).

Research Objective

The primary purpose of this research project was to obtain data that would facilitate the development, validation, and standardization of a measure for DLISt7. As a rule, a measure is said to be standardized when; (a) its rules of measurement are clear, (b) it is practical to apply, (c) it is not demanding of the administrator or respondent, and (d) its results do not depend upon the administrator (Netemeyer, Bearden & Sharma, 2003; Nunnally & Bernstein, 1994). Consequently, a measure that successfully fulfils these criteria would yield "similar results across applications (i.e., the measure is reliable), and offer scores that can be easily interpreted as low, medium [or] high" (Netemeyer et al., 2003, p. 2).

This research project also attempted to ascertain the validity of DLISt7 as a conceptual framework and the reliability of the scale. This was achieved by systematically determining the relationships for the following research questions using accumulated and integrated evidence (Cronbach, 1990). Firstly, how many principles from DLISt7 would actually load significantly? Secondly, would the factor loadings indicate construct validity? Lastly, would an assessment of the summated gain scores reveal the perceived effectiveness of DLISt7 (Dunn-Rankin, Knezek, Wallace & Zhang, 2004; Tuckman, 1999)?

Research Methodology

The study was designed as a non-equivalent pretest-posttest control group Internet quasi-experiment that would "provide substantially better control of the threats to validity than do pre-experimental designs" (Tuckman, 1999, p. 167). Such a design may be used "where better designs are not feasible" (Campbell & Stanley, 1963, p. 204) because "conditions complicate or prevent complete experimental control" (Tuckman, 1999, p. 168).

DLISt7 as a "treatment is included by selection rather than manipulation" (Tuckman, 1999, p. 181) and because of its inherent qualities can also be used as an unobtrusive measure that does not "require acceptance or awareness by the experimental subjects" (Tuckman & Harper, 2012, p. 126). Further replication of the research project and the utilization of the research instrument by others in the future to cross-validate DLISt7 would be valuable in unlocking its Rubik's cube-like potential (Syaril Izwann Jabar, 2012b).

In light of this research being conducted over the Internet, it also qualifies as a field experiment because the research was in a real-life setting (Christensen, 1997, p. 93). The significance of in-the-field Internet experimentation cannot be overlooked because it is useful in terms of determining if a manipulation would work in the real-world (Johnson & Christensen, 2012, p. 285). Moreover, there are also the value-added advantages of speed, low cost, external validity, experimenting around the clock, a high degree of automation of the experiment (i.e., low maintenance, limited experimenter effects), and a wider sample (Reips, 2002, p. 244). However, higher than usual dropout rates [also known as differential attrition (Johnson & Christensen, 2008), differential loss or experimental mortality (Campbell & Stanley, 1963)] are a disadvantage of voluntary participation in Web experiments (Reips, 2000).

Research Sampling

Sample members were drawn using a three-stage purposive cluster sampling technique (Ary, Jacobs & Sorenson, 2010; Cochran, 1977; Johnson & Christensen, 2008). The first sampling element used was of nationality. This was followed by the second element of how far the participants had progressed in their degrees. The third sampling element was of academic affiliation across five faculties at a regional Australian multi-campus university. Participants were recruited based on enrolment in intact courses subject to approval from Faculty. Full ethics clearance was granted by the university's fast track Human Research Ethics Committee (H10REA016). The whole process of research sampling took sixteen months to complete before any data could be collected.

Research Instrument

Reliability Analysis

In an effort to answer the not-so-simple question to which "there are legitimate disagreements about the correct answer" regarding the issue of how are the measurement of constructs developed and validated (Nunnally & Bernstein, 1994, p. 86) the "psychometrical properties of the questionnaire, such as construct validity and reliability" are discussed in the following section (Vandewaetere & Desmet, 2009, p. 349). In extending on a previous investigation which replicated Guidera's (2003) doctoral research project at the masters' level (2004-2007) a variant of Ehrmann, Gamson, and Barsi's (1989) Faculty Inventory was translated, rephrased, and adapted for use as a Student Inventory.

The objectivity and content validity of the adapted version of the research instrument was informally evaluated by a panel of experts consisting of one Associate Dean of Academic Affairs and two subject matter experts (SME), one from the Faculty of Education and Human Development, and the other from the Faculty of Languages and Communication at a teacher training university in Malaysia. The instrument included multiple items for each of the Seven Principles. Each item was presented as a statement eliciting students' perception about the degree of success to which instructional strategies were being effectively or ineffectively used to conduct online learning using a five (5)-point Likert scale.

For the pilot study an excellent value for Cronbach's alpha ($\alpha = 0.97$, $n = 74$) was obtained with individual items having alphas ranging from a lower limit (LL) of 0.972 to an upper limit (UL) of 0.974. For the main study a slightly lower but still excellent value for alpha ($\alpha = 0.94$, $N = 397$) was obtained with individual items having alphas ranging from 0.938 (LL) to 0.941 (UL) (George & Mallery, 2011; Syaril Izwann Jabar, 2007). No problematical items were identified requiring exclusion from the instrument (Coakes & Ong, 2011).

This was followed by an exploratory factor analysis (EFA) to determine the construct validity of the intangible constructs that constitute the Seven Principles. A principal component analysis (PCA) was conducted on the 34 items to verify construct validity. The rotated component matrix revealed that of the 34 items used, 23 were pure variables, while 11 were complex variables (Coakes, Steed & Price, 2008). However, these complex variables did not have loadings that made their structure ambiguous and interpretation difficult (Syaril Izwann Jabar, 2007).

More recent revisions to the measurement instrument being developed at the postgraduate level (2009-2013) involved attaching DLIS to the Seven Principles framework to form DLISt7. A set of four items addressing elements of DLIS were added to the beginning of the instrument addressing each of the four levels (0-3) from presentation of Information only (Level 0) to Task-centred integration (Level 3). The sets of items for each principle were then overlayed with attributes associated with students' perception. The Likert scales were also switched to a Sentence Completion Rating scale "with descriptive statements on either end" (Tuckman & Harper, 2012, p. 229).

This adjustment was made to sidestep "the multidimensionality innate in Likert-type scales" and eliminate "the extra cognitive load associated with the use of item reversals" (Hodge, 2007, p. 289). Furthermore, the use of such a scale would be an improvement in terms of fulfilling parametric assumptions and coping with issues such as "coarse response categories" and "equating the neutral option with a not applicable response" (Hodge & Gillespie, 2003, p. 53).

Accordingly, the perception of the participants towards DLISt7 was successfully measured using scores that can be easily interpreted as low, medium or high (Netemeyer et al., 2003). Cronbach's alpha for the pilot study revealed an excellent coefficient ($\alpha = 0.92$, $n = 39$) with results indicating that removal of individual items would result in alphas ranging from 0.913 (LL) to 0.918 (UL). In the main study, a slightly higher alpha coefficient ($\alpha = 0.95$, $n = 283$) was obtained using a larger sample with removal of individual items resulting in alphas ranging from 0.950 (LL) to 0.952 (UL). Again, there were no problematical items identified (Coakes & Ong, 2011) and by assessing alpha coefficients from both the pilot and main study it was determined that the temporal stability of the measure was excellent (George & Mallery, 2011).

Thus, the internal consistency of the measurement instrument has held up well because when estimating the "correlation (reliability coefficient) to be expected if two independent, more or less equivalent forms of a test are applied on the same occasion," it is expected that "the stronger the intercorrelations among a test's items, the greater its homogeneity" (Cronbach, 1990, p. 704). Although validation can be obtained from a single study, "the ideal is a process that accumulates and integrates evidence on appropriateness of content, correlations with external variables, and hypotheses about constructs" (Cronbach, 1990, p. 707).

Results

Students' Awareness of DLISt7

The questionnaire (Appendix) was administered online using LimeSurvey® (limesurvey.com). Faculty members responsible for selected courses at a regional Australian university cooperated by including the universal resource locator (URL) for the survey in messages to students. A total of 319 completed responses were collected for the pre-test administered at the start of the semester.

Regardless of gender, nationality, academic progress, or faculty affiliation, 194 (60.8%) indicated "yes" they were aware of DLISt7 and 125 (40.0%) indicated "no," they were not aware. In view of DLISt7 being an unpublished conceptual framework, it is doubtful that undergraduate students from this university could have had a priori knowledge about it. Any claim made contrary to the fact could have occurred purely by chance, but is more likely to have been a combination of circumstances that cannot be isolated without further study. The following reasons are suggested to explain why there was a sixty/forty split in responses.

Firstly, it is possible that it was a case of the Hawthorne effect, which "refers to performance increments prompted by mere inclusion in an experiment" (Tuckman & Harper, 2012, p. 132). This could have come to pass because of a fumble when allocating participants to the No Treatment-Treatment conditions that might have tipped-off or aroused the suspicion of a few participants (McMillan & Schumacher, 2009).

Secondly, it is also possible that performance on the posttest was affected by experience from the pre-test (Tuckman & Harper, 2012). Problems related to testing occur because "experience of taking such a pre-test may increase the likelihood that the subjects will improve their performance on the subsequent posttest, particularly when it is identical to the pre-test" (Tuckman & Harper, 2012, p. 126).

Thirdly, is the predisposition to "provide the answer they want others to hear about themselves rather than the truth.... that shows oneself in the best possible light," known as the social desirability response bias (Tuckman & Harper, 2012, p. 265).

Lastly, there is also the "tendency to mark a single choice for all questions out of boredom, disinterest, or hostility" (Tuckman & Harper, 2012, p. 264) known as "acquiescent" response bias (Cronbach, 1990, p. 470).

The Utilization of Communication Technology and Online Resources by Teaching Staff

The most frequently utilized communication technology or online resource used for conveying instructional strategies was StudyDesk with 284 (89.0%) "yes" responses followed by email with 256 (80.3%) "yes" responses. This was followed in descending order by Wimba Online Classrooms (f = 82, 25.7%), Moodle Forums (f = 81, 25.4%), blogs (f = 71, 22.3%), telephone: voice (f = 32, 10.0%), Moodle Chat (f = 29, 9.1%), and instant messaging (f = 28, 8.8%). Hence, it would probably be reasonable to assume that teaching staff at this university had the tendency to rely heavily on two of the more important communication technologies (StudyDesk and email) made available to them while preferring to be parsimonious when choosing what other online resources to incorporate into their instructional repertoire. Possible reasons include lack of familiarity with less common resources or the desire to not overwhelm the students.

The Interaction between Awareness of DLISt7, No Treatment-Treatment Group, and Gender

Initial findings revealed that students' Awareness of DLISt7 at the pretest stage was independent of or not related to being in the No Treatment-Treatment group. However, students' Awareness of DLISt7 at the posttest stage was related to being in the No Treatment-Treatment group. Thus, the need arose to further investigate whether it would be probable to assume that the intervening variable was actually transmitting or mediating the effect of the treatment variable onto the dependent variable (Creswell, 2012), or was it a case of uncontrolled extraneous variables confounding the results, and "casting doubt about the validity of inferences made" (Pedhazur & Schmelkin, 1991, p. 212).

A higher order between-subjects three-way ANOVA revealed that there was no statistically significant interaction between the posttest scores for Awareness, No Treatment-Treatment group, and Gender. However, there was a statistically significant main effect for the No Treatment-Treatment group. It was also established that students' Awareness of DLISt7 at the posttest stage was related to being in the No Treatment-Treatment group and that pre- and posttest scores were related.

Once again the researcher had to attempt to comprehend the source and nature of why the mean scores for the posttest were not significantly greater than the mean scores for the pretest (Cohen, Cohen, West & Aiken, 2003). From where did the uncertainty overshadowing "the accuracy of the inferences, interpretations, or actions made on the basis of test scores" provided by the participants from the No Treatment group originate (Johnson & Christensen, 2012, p. 597)? The following explanation was proposed to clarify how the confounding variable of group mean scores and the extraneous factor of group sample size for the No Treatment group have come together to limit the reliability and validity of the inferences derived from the findings.

In view of DLISt7 being an unpublished conceptual framework, it is doubtful that undergraduate students from the No Treatment group, could have had a priori knowledge about DLISt7. Although the probability does exist that such responses could have occurred purely by chance, logic favours the assumption that the responses were confounded, either by the Hawthorn or testing effect, together with the social desirability and acquiescent response bias. As stated by McMillan and Schumacher (2009), "scores cannot be valid unless they are reliable…. Reliability is needed for validity; scores can be reliable but not valid" (p. 185). Hence, the only way to know for sure is to conduct further research using the Solomon four-group design (McMillan & Schumacher, 2009).

Justification for this alternative explanation was realized while conducting a detailed analysis of the mean scores for the No Treatment-Treatment group. For the pretest, the mean score for the No Treatment group was 79.85 (N = 63) but, for the posttest the mean score was 84.23 (N = 34). It was also determined that the pretest mean score for the Treatment group was 76.37 (N = 220) and the posttest mean score was 76.26 (N = 82), an insignificant difference that did not register on any of the statistical tests but could be further investigated.

The primary point of contention that warranted careful consideration was whether the posttest mean scores for the No Treatment-Treatment group (M=84.23) were representative of the population mean since they were from a sample of 34 participants made up of 29 Females (M = 85.25) and 5 Males (M = 78.26). Consequently, the posttest mean scores from the No Treatment group would appear inflated when compared to the posttest mean scores from the Treatment group (M = 76.26) which was from a larger sample of 82 participants made up of 53 Females (M = 79.18) and 29 Males (M = 70.93). As a result, the mean scores that came from the latter Treatment group and not the former No Treatment group would appear to best represent the population mean without giving the impression of being overstated.

The next point of contention is the fact that a sample of approximately 40 would have been better for invoking the central limit theorem (Field, 2009). With a sample of less than 30, the resulting sampling distribution would have a different shape compared to the parent population causing doubt about whether "the sampling distribution has a normal distribution with a mean equal to the population mean" (Field, 2009, p. 42).

According to Glass and Hopkins (1996), "the validity of the central limit theorem allows [for] statistical inferences to [be made across] a much broader range of applications than would otherwise be possible" (p. 235). This theorem applies "even when the parent population is not normal, [because] the formula $\sigma_{\bar{x}} = \sigma / \sqrt{n}$ accurately depicts the degree of variability in the sampling distribution" (Glass & Hopkins, 1996, p. 239).

For example, when "sample sizes are small (1, 2, 5 and 10); some degree of non-normality in the parent population continues to be evident in the sampling distributions, but progressively less so as *n* increases" (Glass & Hopkins, 1996, p. 239). When *n* was increased to 25, the theoretical standard error of the mean agrees almost perfectly with the standard deviation from the sample means despite a skewed parent population from which the sample was drawn (Glass & Hopkins, 1996).

The Validity of DLISt7 and the Reliability of Its Items

Findings from the exploratory and confirmatory factor analysis verified the validity of the intangible constructs that constitute the conceptual framework of DLISt7. Even though only seven of the possible eight factors submitted successfully loaded, there is a simple and logical explanation. According to Brown (2006), the number of observed measures (p) that are submitted for analysis limits the number of factors that can be extracted (m). Unequivocally, $p - 1$ is the maximum number of factors that can be extracted or "the number of parameters that are estimated in the factor solution (a) must be equal to or less than the number of elements (b) in the input correlation or covariance matrix (i.e., $a \leq b$)" (Brown, 2006, p. 23).

The factor correlation matrix revealed an uncorrelated model with the oblique rotation producing a solution that was "virtually the same as one produced by orthogonal rotation" (Brown, 2006, p. 32). In fact, the interpretation of the oblique solution, although more complicated than the orthogonal solution, did provide results that were better (Hatcher, 2007). Together with the fact that the test-retest (temporal) coefficient for Cronbach's alpha reliability analysis was excellent each time the research instrument was administered, it would probably be safe to assume that the items are actually measuring "the underlying construct comparably across groups" (Brown, 2006, p. 4).

Discussion

Validation as a process is unending and requires measures to be "constantly evaluated and re-evaluated to see if they are behaving as they should" (Nunnally & Bernstein, 1994, p. 84). As assured by Cronbach (1990), "the more reliable a measuring procedure is, the greater the agreement between scores obtained when the procedure is applied twice" (p. 705). Thus, there would be value in conducting further research using the revised version of the measurement instrument. Not only would replication make available a fresh reliability index based on test-retest reliability, but factor analysis could then be used to refine the precise measurement of constructs and contingencies, or in other words, the convergent or discriminate validity of DLISt7 (Netemeyer et al., 2003; Rust & Golombok, 1989; Skinner, 1954).

Hence, the real world utility of factor analysis is the ability to "summarize the interrelationships among the variables in a concise but accurate manner as an aid in conceptualization" (Gorsuch, 1983, p. 2). A conceptual framework is only as good as it can "reduce the amount of trial-and-error effort, and people who explore theories stand at the vanguard of each field of science" (Nunnally & Bernstein, 1994, p. 317). Only through such efforts can the hierarchical levels of a construct, also known as depth psychometry, be studied (Cattell & Schuerger, 1978, p. 223).

Conclusion

Based on Chickering and Gamson's Seven Principles, this study attempted to revitalize the principles by merging them with Merrill's DLIS. The goal was to develop, validate, and standardize a measure for assessing the effectiveness of DLISt7. As a measurement instrument, DLISt7 has been successfully standardized because; (a) its rules of measurement are clear, (b) it is practical to apply, (c) it is not demanding of the administrator or respondent, and (d) its results do not depend upon the administrator (Netemeyer et al., 2003; Nunnally & Bernstein, 1994). Consequently, DLISt7 successfully fulfils all the relevant criteria and has yielded similar scores across applications that can be easily interpreted as low, medium or high, indicating that as a measurement model it is reliable (Netemeyer et al., 2003).

Inductively, the research questions were also successfully answered using a systematic approach. Firstly, of the eight principles specified, seven loaded successfully. Secondly, from the factor loadings it was ascertained that the items utilized are measuring the appropriate constructs seeing as accumulated and integrated evidence indicate that such a conclusion would be appropriate (Cronbach, 1990). However, an assessment of the summated gain scores about the perceived effectiveness of DLISt7 was inconclusive.

Furthermore, DLISt7 also meets the terms of Nunnally and Bernstein's (1994) three major aspects of construct validation, namely: (1) the domain of observables related to the construct have been specified, (2) the extent to which the observables tend to measure the same things has been determined, and (3) individual differences studies or experiments that attempt to determine the extent to which the supposed measures of construct are consistent with best guesses have also been performed. The resultant standardized measure can now be used as a rubric to assist the design of instruction and as a measure to evaluate the effectiveness of such instruction.

Thus, it is proposed that DLISt7 be utilized to enable the learning experienced by students to be systematically scalable to different levels of complexity. Teaching staff would conceivably have the flexibility of being eclectic in their choice of pedagogy for providing students with directed facilitation (Shea, Li & Pickett, 2006) to work their way through the pathways of knowledge to find their own answers. Successively less facilitated guidance, also known as guided instruction (Kirschner, Sweller & Clark, 2006) should be faded with each scaffolded task until students are completing complex tasks on their own.

Will any of this make a difference in bridging and connecting with what is there and not having to reinvent the wheel? Anderson and Dron (2011) summed up the idea well when they said that in order to identify the best mix of pedagogy and instructional technology, the learning and teaching experience has to be seen as a progression. Over the past three decades, many technologies have come and gone, and so has the popularity of different approaches to pedagogy. But each has built upon the shortcomings of the instructional technology left behind by its predecessor instead of replacing the first of its kind.

To recall and demonstrate an understanding of how to apply and integrate what use to be referred to as the Socratic method, a teacher has always, and will be expected to continue the tradition of setting good examples by being "clear about his objectives; he knows why he is doing what he is doing; and he chooses a technique to suit his objective" (Hyman, 1974, p. 101).

References

Albion, P. R., & Redmond, P. (2006). *Returning the favour: Using insights from online learning to enhance on-campus courses*. Paper presented at the 17th International Conference of the Society for Information Technology & Teacher Education (SITE 2006), Orlando, U.S.A. Retrieved September, 21, 2012, http:///www.editlib.org/f/22444

Anderson, T., & Dron, J. (2011). Three generations of distance education pedagogy. *The International Review of Research in Open and Distance Learning, 12* (3). Retrieved August 23, 2012, from http://www.irrodl.org/index.php/irrodl/article/view/890/1663

Anderson, T., & Dron, J. (2012). Learning technology through three generations of technology enhanced distance education pedagogy. *European Journal of Open, Distance and E-Learning, 2*: 1-16. Retrieved January 9, 2013, from http://www.eurodl.org/materials/contrib/2012/Anderson_Dron.pdf

Ary, D., Jacobs, L. C., & Sorenson, C. (2010). *Introduction to research in education* (8th ed.). Belmont: Wadsworth.

Bangert, A. W. (2004). The seven principles of good practice: A framework for evaluating on-line teaching. *The Internet and Higher Education, 7*(3), 217-232. Retrieved September 16, 2009, from http://dx.doi.org/10.1016/j.iheduc.2004.06.003

Bangert, A. W. (2008a). The influence of social presence and teaching presence on the quality of online critical inquiry. *Journal of Computing in Higher Education, 20*(1), 34-61. Retrieved May 27, 2011, from https://www.springerlink.com/content/g054232413m24746/fulltext.pdf

Bangert, A. W. (2008b). The development and validation of the student evaluation of online teaching effectiveness. *Computers in the Schools, 25*(1), 25-47. Retrieved September 16, 2009, from http://dx.doi.org/10.1080/07380560802157717

Batts, D. (2008). Comparison of student and instructor perceptions of best practices in online technology courses. *MERLOT Journal of Online Learning and Teaching, 4*(4). Retrieved June 22, 2009, from http://jolt.merlot.org/vol4no4/batts_1208.htm

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: The Guildford Press.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research (on teaching).* Boston: Houghton Mifflin Company.

Case, R., & Bereiter, C. (1984). From behaviourism to cognitive behaviourism to cognitive development: Steps in the evolution of instructional design*. *Instructional Science, 13,* 141-158.

Cattell, R. B., & Schuerger, J. M. (1978). *Personality theory in action*. Champaign: Institute for Personality and Ability Testing (IPAT).

Chickering, A. W., & Gamson, Z. F. (1987). *Seven principles for good practice in undergraduate education.* Retrieved July 26, 2005, from http://honolulu. hawaii.edu/intranet/committees/FacDevCom/guidebook/techtip/7 princip.htm

Chickering, A. W., & Ehrmann, S. C. (1996). *Implementing the seven principles: technology as lever*. Retrieved August 4, 2005, from http://www.tltgroup.org/ programs/seven.html

Chickering, A. W., & Gamson, Z. F. (1999). Development and adaptations of the seven principles for good practice in undergraduate education. *New Directions for Teaching and Learning, 80*: 75-81. Retrieved January 25, 2010, from http://dx.doi.org/10.1002/tl.8006

Christensen, L. B. (1997). *Experimental methodology* (7th ed.). Needham Heights: Allyn & Bacon.

Coakes, S. J., Steed, L., & Price, J. (2008). *SPSS: Analysis without anguish; version 15.0 for windows.* Milton: John Wiley & Sons.

Coakes, S. J., & Ong, C. (2011). *SPSS: Analysis without anguish; version 18.0 for windows.* Milton: John Wiley & Sons.

Coates, H. (2006). *Student engagement in campus-based and online education*. Abingdon: Routledge.

Cobbett, S. (2007). A re-conceptualized model of good online pedagogical practices. In T. Bastiaens & S. Carliner (Eds.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2007* (pp. 2410-2415) by the Association for the Advancement of Computing in Education (AACE). Chesapeake: VA, Retrieved September 16, 2009, from http://www.editlib.org/f/26718

Cochran, G. W. (1977). *Sampling techniques*. (3rd ed.). New York: John Wiley & Sons.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences.* Mahwah: Lawrence Erlbaum, Associates.

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston: Pearson Education.

Cronbach, L. J. (1990). *Essential of psychological testing*. New York: Harper & Row.

Dunn-Rankin, P., Knezek, G. A., Wallace, S., & Zhang, S. (2004). *Scaling Methods* (2nd ed.). Mahwah, New Jersey: Lawrence Erlbaum, Associates.

Field, A. (2009). *Discovering statistics using SPSS* (3rd Ed.). London: Sage Publication.

Garrison, D. R., & Cleveland-Innes, M. (2005). Facilitating cognitive presence in online learning: Interaction is not enough. *American Journal of Distance Education, 19*(3), 133-148. http://dx.doi.org.ezproxy.usq.edu.au/10.1207/s15389286ajde1903_2

Garrison, D. R., Cleveland-Innes, M., & Fung, T. S. (2010). Exploring causal relationships among teaching, cognitive and social presence: Student perceptions of the community of inquiry framework. *The Internet and Higher Education, 13*(1-2), 31-36. http://dx.doi.org.ezproxy.usq.edu.au/10.1016/j.iheduc.2009.10.002

George, D., & Mallery, P. (2011). *SPSS for windows step by step: A simple guide and reference 18.0 update* (11th ed.). Boston: Allyn & Bacon.

Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education and psychology.* Needham Heights: Allyn & Bacon.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). New Jersey: Lawrence Erlbaum, Associates.

Grant, M. R., & Thornton, H. R. (2007). Best practices in undergraduate adult-centered online learning: mechanisms for course design and delivery. *MERLOT Journal of Online Learning and Teaching, 3*(4), 346-356.

Guidera, S. G. (2003). Perceptions of the effectiveness of online instruction in terms of the seven principles of effective undergraduate education. *Journal of Educational Technology Systems, 32*(2 & 3), 139-178.

Hatcher, L. (2007). *A step-by-step approach to using SAS for factor analysis and structural equation modelling* (9th ed.). North Carolina: SAS Publishing.

Haughton, N. A., & Romero, L. (2009) The online educator: Instructional strategies for effective practice. *MERLOT Journal of Online Learning and Teaching, 5*(3), 570-576.

Hodge, D. R. (2007). The spiritual competence scale: A new instrument for assessing spiritual competence at the programmatic level. *Research on Social Work Practice*, *17*(2), 287-295. Retrieved June 2, 2009, from http://rsw.sagepub.com/cgi/content/abstract/17/2/287

Hodge, D. R., & Gillespie, D. F. (2003). Phrase completions: An alternative to likert scales. (Note on Research Methodology). *Social Work Research, 27* (1), 45-55. Retrieved May 27, 2009, from http://www.articlearchives.com/science-technology/behavior-cognition-psychology/986300-1.html

Hyman, T. (1974). *Ways of teaching* (2nd ed.). New York: Harper & Row Publishers.

Johnson, B., & Christensen, L. (2008). *Educational research: Quantitative, qualitative and mixed approaches* (3rd ed.). Thousand Oaks: Sage Publications.

Johnson, B., & Christensen, L. (2012). *Educational research: Quantitative, qualitative and mixed approaches* (4th ed.). Thousand Oaks: Sage Publications.

Kehrwald, B., Reushle, S., Redmond, P., Cleary, K., Albion, P., & Maroulis, J. (2005). Online pedagogical practices in the Faculty of Education at the University of Southern Queensland. *Working Paper Series* (pp. i - 33). Toowoomba, Queensland: University of Southern Queensland.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75-86. Retrieved October 4, 2012 from http://www.tandfonline.com/doi/abs/10.1207/s15326985ep4102_1#tabModule

Kuh, G. D., Pace, C. R., & Vesper, N. (1997). The development of process indicators to estimate student gains associated with good practices in undergraduate education. *Research in Higher Education, 38*(4), 435-454.

Kuh, G. D., & Hu, S. (2001). The effects of student-faculty interaction in the 1990s. *The Review of Higher Education, 24*(3), 309-332.

Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.

Lever-Duffy, J., McDonald, J. B., & Mizell, A. P. (2003). *Teaching and learning with technology.* Boston: Pearson Education.

Lever-Duffy, J., McDonald, J. B., & Mizell, A. P. (2005). *Teaching and learning with technology* (2nd ed.). Boston: Pearson Education.

Lever-Duffy, J. & McDonald, J. B. (2008). *Teaching and learning with technology* (3rd ed.). Boston: Pearson Education.

McMillan, J. H., & Schumacher, S. (2009). *Research in education: Evidence-based inquiry* (7th ed.). New Jersey: Prentice-Hall.

Merrill, M. D. (2006). *Hypothesized performance on complex tasks as a function of scaled instructional strategies*. 1-33. Retrieved October 26, 2009, from http://cito.byuh.edu/merrill/text/papers/Scaled%20Strategies.pdf

Merrill, M. D. (2008). What makes e^3 (effective, efficient, engaging) instruction? *Paper presented at the Association for Education Communication and Technology (AECT) Research Symposium.* Retrieved March 23, 2010 from www.aect.org/events/symposia/Docs/Merrill_paper.doc

Merrill, M. D. (2009). First principle of instruction. In C.M. Reigeluth & A. A. Carr-Chellman (Eds.), *Instructional design theories and models: Building a common knowledge base* (Vol. III, pp. 41-56). Madison Ave: Routledge.

Morrison, R. G., Ross, M. S., & Kemp, E. J. (2001). *Designing effective instruction* (3rd ed.). New York: John Wiley & Sons.

Netemeyer, R. G., Beardon, W. O., & Sharma, S. (2003). *Scaling procedures: Issues and applications*. Thousand Oaks: Sage Publications.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Pedhazur, E. J., & Schmelkin, P. (1991). *Measurement, design, and analysis: An integrated approach* (Student ed.). New Jersey: Lawrence Erlbaum, Associates

Reips, U.-D. (2000). The web experiment method: Advantages, disadvantages, and solutions. In M. H. Birnbaum, *Psychological experiments on the internet* (pp. 89-114). San Diego: Academic Press.

Reips, U.-D. (2002). Standards for internet-based experimenting. *Experimental Psychology, 49*(4), 243-256.

Reiser, R. A. (2012). What field did you say you were in? Defining and naming our field. In R. A. Reiser & J. V. Dempsey (Eds.), *Trends and issues in instructional design and technology* (3rd ed., pp. 1-7). Boston: Allyn & Bacon.

Rust, J., & Golombok, S. (1989). *Modern psychometrics.* London: Routledge.

Scardamalia, M., & C. Bereiter. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 97-118). New York: Cambridge University Press**.** Retrieved September 24, 2009, from http://ikit.org/fulltext/2006_KBTheory.pdf

Shea, P., Li, C. S., & Pickett, A. (2006). A study of teaching presence and student sense of learning community in fully online and web-enhanced college courses. *The Internet and Higher Education,* 9, 175−190.

Shea, P., Hayes, S., Smith, S. U., Vickers, J., Bidjerano, T., Pickett, A., Gozza-Cohen, M., Wilde, J., & Jian, S. (2012). Learning presence: Additional research on a new conceptual element within the community of inquiry (CoI) framework. *The Internet and Higher Education, 15*(2), 89-95.

Shea, P., Hayes, S., Smith, S. U., Vickers, J., Bidjerano, T., Gozza-Cohen, M., Jian, S., Pickett, A., Wilde, J., & Tseng, C. (2013). Online learner self-regulation: Learning presence viewed through quantitative content-and social network analysis. *The International Review of Research in Open and Distance Learning, 14*(3), 427-461.

Shneiderman, B. (1998). Relate-create-donate: A teaching/learning philosophy for the cyber-generation. *Computers & Education, 31*, 25-39.

Skinner, B. F. (1954). The science of learning and the art of teaching. *Harvard Educational Review, 24*, 86-97.

Skinner, B. F. (1962). *Walden two* (Paperback ed.). New York: The Macmillan Company.

Skinner, B. F. (1968). *The technology of teaching*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Spector, J. M., Ohradza, C., Van Schaack, A., & Wiley, D. A. (2005). *Innovations in instructional technology; Essays in honor of M. David Merrill.* Mahwah: Lawrence Erlbaum, Associates.

Syaril Izwann Jabar. (2007). *Student perception towards the effectiveness of instructional strategies for online learning in Sultan Idris University of Education (UPSI).* Unpublished M. Ed dissertation. Tanjong Malim: UPSI.

Syaril Izwann Jabar. (2012a). *Assessing the effectiveness of the different levels of instructional strategies [DLISt] for online learning at the University of Southern Queensland (USQ), Australia*. Paper presented at the 2nd Malaysian Postgraduate Conference 2012, Bond University, Gold Coast, Queensland, Australia, http://msda.org.au/index.php/students-corner/malaysian-postgraduate-conference-2012-proceedings

Syaril Izwann Jabar. (2012b). *Assessing the effectiveness of the different levels of instructional strategies [DLISt] for online learning at the University of Southern Queensland (USQ), Australia*. Paper presented at the International Conference on Quality of Teaching and Learning (ICQTL 2012) organized by the Centre of Quality and Academic Development, University Malaysia Terengganu.

Syaril Izwann Jabar. (2012c). *Assessing the effectiveness of the different levels of instructional strategies [DLISt7] for online learning by undergraduate students from the University of Southern Queensland (USQ), Australia*. Unpublished Ph. D dissertation. Toowoomba: USQ.

Tuckman, B. W. (1999). *Conducting educational research* (5th ed.). Orlando: Harcourt Brace College Publishers.

Tuckman, B. W., & Harper, B. E. (2012). *Conducting educational research* (6th ed.). Lanham: Rowman & Littlefield Publishers, Inc.

Vandewaetere, M., & Desmet, P. (2009). Introducing psychometrical validation of questionnaires in CALL research: The case of measuring attitude towards CALL. *Computer Assisted Language Learning, 22*(4), 349-380.

Wuensch, K. L., Shahnaz, A., Ozan, E., Kishore, M., & Tabrizi M. H. N. (2009). Technology and pedagogy: The association between students' perceptions of the quality of online courses and the technologies employed. *MERLOT Journal of Online Learning and Teaching, 5*(2), 253-262.

Appendix: Research Instrument; Online version, 28 July 2010

ASSESSING THE EFFECTIVENESS OF THE DIFFERENT LEVELS OF INSTRUCTIONAL STRATEGIES (DLISt7) FOR ONLINE LEARNING BY UNDERGRADUATE STUDENTS AT THE UNIVERSITY OF SOUTHERN QUEENSLAND (USQ), AUSTRALIA

This survey is designed to measure and assess various attributes associated with students' perception towards the effectiveness of the different levels of instructional strategies that can be used to conduct online learning.

All information provided will remain confidential and your identity will not be disclosed. By completing this survey you are consenting to being a participant in this research.

Please **check the appropriate boxes** or **fill in the blanks** to indicate your response.

a) Are you aware that there are Different Levels of Instructional Strategies that can be used to conduct online learning?
- ☐ Yes
- ☐ No

b) What is your Gender?
- ☐ Female
- ☐ Male

c) What is your Nationality?

____

d) How far have you progressed in your degree at USQ?

Degree : ____

Year (1st, 2nd, etc.,) : ____________________

Semester : ____________________

e) Which Faculty are you from?
- ☐ Arts
- ☐ Business
- ☐ Sciences
- ☐ Education
- ☐ Engineering & Surveying

f) Please check the boxes that indicate the communication technology or online resource utilized by teaching staff to convey instructional strategies for online learning. Check any that apply.

☐ Blogs
☐ Email
☐ StudyDesk
☐ Moodle Chat
☐ Moodle Forum
☐ Teleconferencing
☐ Videoconferencing
☐ Instant Messaging
☐ Wimba Online Classroom
☐ Telephone: Text Messaging
☐ Telephone: Voice
☐ Skype Video
☐ Skype Voice
☐ Skype Text
☐ None are utilized
Other ____________________

The following statements use a sentence completion format to measure various attributes associated with students' perception towards the effectiveness of the different levels of instructional strategies for online learning.

A partially completed sentence is provided, followed by a scale ranging from 1 to 10. The 1 to 10 range provides you with a continuum on which to reply, with *1* corresponding to a *minimum* amount of the attribute, while *10* corresponds to the *maximum* amount of the attribute. A *5* corresponds to an *average* amount of the attribute.

Please **select** a number along the continuum that best reflects your initial feeling.

1.	**DIFFERENT LEVELS OF INSTRUCTIONAL STRATEGY**	**FACTOR LOADING**	**FACTOR RATING**
1.1	I __________ noticed instances of Teaching staff trying to present information with accompanying recall questions. **Rarely** **Frequently** **1** 2 3 4 5 6 7 8 9 **10**	0.570	Very Good
1.2	I recall attempts by Teaching staff to promote higher levels of performance on complex tasks by way of presenting information, and demonstrating its application as being __________. **Meaningless** **Significant** **1** 2 3 4 5 6 7 8 9 **10**	0.669	Excellent
1.3	I can __________ understand why Teaching Staff would be willing to provide corrective feedback in order to promote improvement in my performance on complex tasks. **Completely** **Vaguely** **10** 9 8 7 6 5 4 3 2 **1**	0.481	Good
1.4	I __________ value attempts by Teaching staff to use a task-	0.577	Very good

	centered approach to promote efficiency, effectiveness and engagement. **Scarcely** **Very much** **1** 2 3 4 5 6 7 8 9 **10**		
2.	**ENCOURAGING INTERACTION BETWEEN STUDENTS AND TEACHING STAFF**	**FACTOR LOADING**	**FACTOR RATING**
2.1	I __________ noticed instances of Teaching staff trying to communicate with me. **Rarely** **Frequently** **1** 2 3 4 5 6 7 8 9 **10**	0.812	Overlap
2.2	I recall attempts by Teaching Staff to facilitate informal interaction with me as being __________. **Meaningless** **Significant** **1** 2 3 4 5 6 7 8 9 **10**	0.658	Excellent
2.3	I can __________ understand why Teaching staff would demonstrate a willingness to serve as a mentor to me. **Completely** **Vaguely** **10** 9 8 7 6 5 4 3 2 **1**	0.666	Excellent
2.4	I __________ value attempts by Teaching staff to contact me when I have fallen behind to discuss my study habits, schedules, and other commitments. **Scarcely** **Very much** **1** 2 3 4 5 6 7 8 9 **10**	0.528	Good
2.5	I am __________ of teaching staff attempting to provide extra material or exercises if I lack the essential background knowledge or skills. **Appreciative** **Unappreciative** **10** 9 8 7 6 5 4 3 2 **1**	0.648	Very good
3.	**DEVELOPING RECIPROCITY AND COOPERATION AMONG STUDENTS**	**FACTOR LOADING**	**FACTOR RATING**
3.1	I __________ noticed instances of Teaching staff trying to encourage me to participate in online activities. **Rarely** **Frequently** **1** 2 3 4 5 6 7 8 9 **10**	0.567	Very good

3.2	I recall attempts by Teaching staff to get me to explain difficult ideas or concepts to others within an online learning group as being __________. **Meaningless** **Significant** **1** 2 3 4 5 6 7 8 9 **10**	0.742	Excellent
3.3	I can __________ understand why Teaching staff would demonstrate an eagerness to get me to discuss openly with colleagues through a forum about interests and backgrounds. **Completely** **Vaguely** **10** 9 8 7 6 5 4 3 2 **1**	0.591	Very good
3.4	I __________ value attempts by Teaching staff to utilize a Learning Management System such as *USQStudyDesk* to encourage learning communities in my course. **Scarcely** **Very much** **1** 2 3 4 5 6 7 8 9 **10**	0.667	Excellent
3.5	I am __________ of Teaching staff attempting to get me and my colleagues to work on projects together. **Appreciative** **Unappreciative** **10** 9 8 7 6 5 4 3 2 **1**	0.480	Good

4.	**ENCOURAGING ACTIVE, CONTEXTUAL AND MEANINGFUL LEARNING**	**FACTOR LOADING**	**FACTOR RATING**
4.1	I __________ noticed instances of Teaching staff trying to get me to apply meaningful learning by relating events that happened in real life to what was being learnt. **Rarely** **Frequently** **1** 2 3 4 5 6 7 8 9 **10**	0.640	Very good
4.2	I recall attempts by Teaching staff to get me to apply contextual learning by analyzing real-life contexts as being __________. **Meaningless** **Significant** **1** 2 3 4 5 6 7 8 9 **10**	0.598	Very good
4.3	I can __________ understand why Teaching staff would demonstrate a willingness to link me with professionals who are experts in the field of study so that opinions and ideas can be exchanged. **Completely** **Vaguely** **10** 9 8 7 6 5 4 3 2 **1**	0.696	Excellent
4.4	I __________ value attempts by Teaching staff to encourage me to express myself when I do not understand a particular subject matter. **Scarcely** **Very much** **1** 2 3 4 5 6 7 8 9 **10**	0.705	Excellent

4.5	I am __________ of attempts to include independent study assignments where I seek out information from the Internet and later discuss with Teaching staff the validity of the information and the reliability of its source. **Appreciative** **Unappreciative** **10** 9 8 7 6 5 4 3 2 **1**	0.578	Very good
5.	**GIVING PROMPT FEEDBACK**	**FACTOR LOADING**	**FACTOR RATING**
5.1	I __________ noticed instances of Teaching staff trying to adjust their instructional strategy to include problem solving and task-centered activities that provided me with immediate feedback. **Rarely** **Frequently** **1** 2 3 4 5 6 7 8 9 **10**	0.799	Excellent
5.2	I recall attempts by Teaching staff to provide corrective feedback regarding my performance on problem solving and task-centered activities as being __________. **Meaningless** **Significant** **1** 2 3 4 5 6 7 8 9 **10**	0.740	Excellent
5.3	I can __________ understand why Teaching staff would demonstrate a willingness to politely inquire about my strengths and weaknesses in tutorials, quizzes and tests. **Completely** **Vaguely** **10** 9 8 7 6 5 4 3 2 **1**	0.734	Excellent
5.4	I __________ value attempts by Teaching staff to get me to go online and contact them to discuss my academic progress. **Scarcely** **Very much** **1** 2 3 4 5 6 7 8 9 **10**	0.614	Very good
5.5	I am __________ of attempts by Teaching staff to provide me with an evaluation of my proficiency. **Appreciative** **Unappreciative** **10** 9 8 7 6 5 4 3 2 **1**	0.732	Excellent
6.	**EMPHASIZING TIME ON TASK**	**FACTOR LOADING**	**FACTOR RATING**
6.1	I __________ noticed instances of Teaching staff tying to communicate to me that I am expected to complete my assignments promptly. **Rarely** **Frequently** **1** 2 3 4 5 6 7 8 9 **10**	0.589	Very good

6.2	I recall attempts by Teaching staff to deliver course materials, quizzes and assignments online as being __________. **Meaningless** **Significant** **1** 2 3 4 5 6 7 8 9 **10**	0.638	Very good
6.3	I can __________ understand why Teaching staff would demonstrate an eagerness to emphasize to me the importance of diligence, sound self-pacing and scheduling. **Completely** **Vaguely** **10** 9 8 7 6 5 4 3 2 **1**	0.670	Excellent
6.4	I __________ value attempts by Teaching staff to make it clear to me the amount of time that is required to understand complex material. **Scarcely** **Very much** **1** 2 3 4 5 6 7 8 9 **10**	0.723	Excellent

7.	**COMMUNICATING HIGH EXPECTATIONS**	**FACTOR LOADING**	**FACTOR RATING**
7.1	I __________ noticed instances of Teaching staff trying to communicate to me that I am expected to work hard. **Rarely** **Frequently** **1** 2 3 4 5 6 7 8 9 **10**	0.610	Very good
7.2	I recall attempts by Teaching staff to emphasize the importance of holding on to high standards for academic achievement as being __________. **Meaningless** **Significant** **1** 2 3 4 5 6 7 8 9 **10**	0.804	Overlap
7.3	I can __________ understand why Teaching staff would demonstrate a willingness to share with me past experiences, attitudes and values. **Completely** **Vaguely** **10** 9 8 7 6 5 4 3 2 **1**	0.683	Excellent
7.4	I __________ value attempts by Teaching staff to provide me with a pre-test at the beginning of the course. **Scarcely** **Very much** **1** 2 3 4 5 6 7 8 9 **10**	0.540	Good
7.5	I am __________ of attempts by Teaching staff to discuss my academic progress especially near the end of the course. **Appreciative** **Unappreciative** **10** 9 8 7 6 5 4 3 2 **1**	0.717	Excellent

8.	RESPECTING DIVERSE TALENTS AND WAYS OF LEARNING	FACTOR LOADING	FACTOR RATING
8.1	I __________ noticed instances of Teaching staff trying to determine my learning style, interests or background at the beginning of the course. **Rarely** **Frequently** **1** 2 3 4 5 6 7 8 9 **10**	0.878	Overlap
8.2	I recall attempts by Teaching staff to relate learning activities to my learning style, interests or background as being __________. **Meaningless** **Significant** **1** 2 3 4 5 6 7 8 9 **10**	0.796	Excellent
8.3	I can __________ understand why Teaching staff would demonstrate a willingness to use multiple methods to communicate their own expectations at the beginning of the course. **Completely** **Vaguely** **10** 9 8 7 6 5 4 3 2 **1**	0.652	Excellent
8.4	I __________ value attempts by Teaching staff to encourage mastery learning or learning contracts as instructional strategies. **Scarcely** **Very much** **1** 2 3 4 5 6 7 8 9 **10**	0.576	Very good
8.5	I am __________ of attempts by Teaching staff to work with me to set challenging objectives for learning outcomes. **Appreciative** **Unappreciative** **10** 9 8 7 6 5 4 3 2 **1**	0.637	Very good

g) With regards to the effectiveness of the proposed Different Levels of Instructional Strategies for Online Learning, I would like to suggest that;

h) This is because I believe that such strategies would;

Thank you for completing this survey

Retention, Progression and the Taking of Online Courses

Scott James
Predictive Analytics Reporting Framework

Karen Swan
University of Illinois Springfield

Cassandra Daston
Predictive Analytics Reporting Framework

Abstract

Online learning continues to grow at post-secondary institutions across the United States, but many question its efficacy, especially for students most at-risk for failure. This paper engages that issue. It examines recent research on the success of community college students who take online classes and explores similar comparisons using 656,258 student records collected through the Predictive Analytics Reporting (PAR) Framework. In particular, the research investigated retention rates for students in three delivery mode groups – students taking only onground courses, students taking only online courses, and students taking some courses onground and some courses online at five primarily onground community colleges, five primarily onground four-year universities, and four primarily online institutions.

Results revealed that taking some online courses did not result in lower retention rates for students enrolled in primarily onground community colleges participating in the PAR Framework. Moreover, although retention rates were lower for such students taking only online courses than for similar students taking only onground or blending their courses, much of the difference could be explained by extraneous factors. Essentially no differences in retention between delivery mode groups were found for students enrolled in primarily onground four-year universities participating in the PAR Framework, while at participating primarily online institutions, students blending their courses had slightly better odds of being retained than students taking exclusively onground or exclusively online courses. No differences between the latter groups were found at these institutions.

Patterns of retention were similar regardless of gender across institutional categories, and were mostly similar regardless of Pell grant status with the exception of fully online students at traditional community colleges. Age, however, did differentially affect delivery mode effects. Older students taking only online courses were retained at higher rates than younger students taking only online courses at both primarily onground community colleges and primarily online institutions. The results suggest that, despite media reports to the contrary, taking online courses is not necessarily harmful to students' chances of being retained, and may provide course-taking opportunities that otherwise might not be available, especially for nontraditional students.

Introduction

Online learning continues to grow at post-secondary institutions across the United States, but many question its efficacy, especially for students most at-risk for failure. This paper engages that issue. It examines recent research on the success of community college students who take online classes and explores similar comparisons using 656,258 student records collected through the Predictive Analytics Reporting (PAR) Framework. In particular, it investigates retention rates for students in three delivery mode groups – students taking only onground courses, students taking only online courses, and students taking some courses onground and some courses online at five primarily onground community colleges, five primarily onground four-year universities, and four primarily online institutions. It also explores potentially differential effects of delivery mode related to Pell grant status, gender, and/or age.

In the sections which follow, relevant research on the effects of online learning on the success of community college students is summarized, and the PAR Framework is explained. The Methodology section identifies the research questions addressed, the data sources used, and the methods of analyses. In the Results section, findings are given for primarily onground community colleges, primarily onground four-year universities, and primarily online institutions broken out by research questions. The Discussion section explores the implications of some of the findings, examines results across institutions, and notes the limitations of the research.

Finally, the major findings of the study are reiterated in the Conclusions.

Background

Online learning is no longer an anomaly in American higher education. According to national data, in the fall 2013 semester over 5.2 million, or 25% of all higher education students in the United States took at least one online course (Allen & Seaman, 2015) and indications are that online learning will continue to grow in the near future.

In addition, most researchers agree that learning outcomes from online courses are not significantly different from traditional courses (Bernard et al., 2009; Means, Toyama, Murphy, Bakia, & Jones, 2009). However, as educators have come to accept the similarity of learning regardless of delivery mode, they have also come to believe that retention and progression are greater problems online (Allen & Seaman, 2015; Moore & Fetzner, 2009). Indeed, there have been several, relatively recent, large-scale studies comparing retention and progression for community college students taking online and traditional classes that support such a view.

Two such studies were undertaken by the Community College Research Center (CCRC) at Teachers College, Columbia University involving cohorts in the Virginia (Jaggers & Xu, 2010) and Washington state (Xu & Jaggers, 2011) community college systems. The 2004 cohorts in both systems were followed for five years. Because the researchers found that better prepared students were more likely to enroll in online courses, they limited their comparisons to the population of students who ever took an online course. They found that in both systems, “ever online” students were more likely to fail or withdraw from online courses than from face-to-face courses (Xu & Jaggers, 2011). In addition, the researchers found that students who took online coursework in early terms were slightly but significantly less likely to return to school in subsequent terms, and students who took a higher proportion of credits online were slightly but significantly less likely to graduate, attain a certificate, or transfer to a four-year institution (Jaggers & Xu, 2010; Xu & Jaggers, 2011).

In a similar, more recent, study of student performance in the cohort enrolling in the California Community College system in the 2008/09 academic year, Hart, Friedman, and Hill (2015) found that students’ likelihood of completing and/or passing courses (receiving a C or better) were lower for online courses than they were for those offered in face-to-face formats. Controlling for possible differences in courses, students, and instructors, the researchers found students were 6.8 to

8.9 percentage points less likely to complete, and 10.9 to 15.2 percentage points less likely to pass online courses.

Shea and Bidjerano (2014), however, reported seemingly different findings using somewhat different outcome measures. Using data from the Beginning Post-secondary Survey (BPS), they compared the degree and/or certificate completion rates of community college students who took one or more courses online with those of students who did not take any online courses. The BPS data comes from a sample of students who initially enrolled in a US post- secondary degree program in 2004. These same students were surveyed again in 2006 and 2009. Shea and Bidjerano explored the data in an attempt to replicate and extend the CCRC findings. They found, however, that controlling for relevant background characteristics, students who took some of their early courses online had a significantly better chance of attaining a community college credential than those who only took face-to-face courses.

Some explanation for these seemingly disparate results can be found in another study of California community colleges. Johnson and Cuellar Majia (2014) studied an earlier cohort who initially enrolled in California community colleges in the fall of 2006. Modeling their work on Xu & Jaggers, the researchers found that students taking online classes were less likely to complete them, and less likely to complete them with a passing grade, than students in enrolled in face-to-face classes. However, when they examined long-term outcomes, Johnson and Cuellar Majia found that students who took at least some online courses were more likely to earn an associate's degree or transfer to a four-year institution than those who didn't.

The research reported in this paper builds on the above studies and investigates the seeming anomalies among them. In particular, it compares both course completion and retention among students enrolled in solely online, solely onground, or both online and onground courses (ever online) across five quite dissimilar primarily onground community colleges. In addition, it tests to see whether or not similar patterns can be found among female vs. male community college students, older vs. younger community college students and/or among community college students receiving or not receiving Pell grants. Moreover, the research also investigates course completion and retention among students enrolled in five quite different four-year colleges and among students enrolled in four very different primarily online institutions. To do so, it uses data collected through the Predictive Analytics Reporting (PAR) Framework.

Predictive Analytics Reporting (PAR) Framework

The PAR Framework is a non-profit, multi-institutional collaborative that provides member institutions with tools and resources for identifying risks and improving student success. PAR member institutions provide anonymized student-level data for all credential-seeking students who began taking courses at the institution in August 2009 or later. At the time of this writing, the PAR data set has more than 2 million student records and 20 million course records from more than 30 institutions and includes data through Fall 2014 for most institutions and through Spring 2015 for some.

These data include

- student demographic information, such as age, gender, race/ethnicity, military and veteran status, permanent residence zip code, Pell eligibility
- prior academic information, including high school GPA, transfer GPA, prior amount and type of college credits earned
- student course information for all courses taken, including specific course titles, course length, course size, outcomes, and delivery mode
- other student academic information, such as majors pursued, specific credentials sought,

transfer credits brought in after enrollment, and credentials earned.

PAR member institutions comprise a range of the many diverse options for post-secondary education, including traditional open admission community colleges, 4-year traditional selective admission public institutions, and nontraditional primarily online institutions, both for-profit and nonprofit.

A key feature of the PAR dataset is the use of PAR's openly published common data definitions by all member institutions. Because all data provided by PAR member institutions utilize these common definitions, cross-institutional "apples to apples" analyses on the combined data set can be performedto better understand the factors that impact student success generally as well as locally.

In addition, having relatively comprehensive, detailed data for all credential-seeking students, rather than a sample from each institution, enables a more accurate understanding of the student and institutional-level factors that impact risk and success. It also makes it possible to more effectively control for confounding variables that might be contributing to observed differences between student groups.

Methodology

The study reported here investigated the effects of delivery mode on the retention and progression of undergraduate students. It explored differences in retention and progression among students who took all their classes online, students who took all their classes onground, and students who blended online and onground classes. The research questions addressed included:

- Do community college students who enroll in online courses have poorer course completion rates and are they retained at lower rates than community college students who take all their courses onground?
- Does delivery mode differentially affect particular groups of community college students?
- Do students enrolled in four-year colleges who take online courses have poorer course completion rates and are they retained at lower rates than four-year college students who take all their courses onground?
- Does delivery mode differentially affect particular groups of four-year college students?
- Are there any differences in course completion and/or retention rates associated with differing delivery modes at primarily online institutions?
- Does delivery mode differentially affect particular groups of students attending primarily online institutions?

Data Sources

In the current study, the impact of course delivery mode on student outcomes in various types of post-secondary settings was explored. Fourteen PAR member institutions were included in the study:

- 5 primarily onground community colleges (213,056 student records)
- 5 primarily onground 4-year universities (113, 036 student records)
- 4 primarily online institutions (330,166 student records)

Data sources were all student- and course-level records for students who began their studies between September 2009 and December 2012 at these schools. Thus, all students included in the analyses had the opportunity for at least 18 months of enrollment in order to determine retention.

Students at each institution were grouped according to delivery mode—fully on- onground, fully online, or a blend of onground and online—based on the courses they enrolled in up to and immediately following the first six months of their enrollment at the institution. Students were considered fully onground if they only took onground courses during that period; they were considered fully online if they only took online courses during that period, and they were considered blended if they took any combination of onground and online courses during that period.

Because post-secondary institutions have course enrollment periods of differing lengths (e.g., semesters, quarters, continuous short- or long-course enrollment periods), the following approach was used to determine a student's course-taking behavior, credits attempted, and credit ratio in their first few months at the institution. For each student included in the study, all courses taken during the student's first six months' enrollment, plus the next course or courses completed after the six-month date, were used to determine delivery mode, credit ratio, and credits attempted. If more than one course ended on the same end date, all were included. If the student stopped taking courses prior to the six-month point, those courses were included. Thus, for all institutions in the study, delivery mode, credit ratio and credits attempted were based on approximately eight to nine months of course data. For traditional semester schools, one academic year was typically included; for quarter schools, three quarters; for continuous enrollment school, eight to nine months of course-taking. Variables regarding credits (including delivery mode, credits attempted, and credit ratio) were measured in aggregate for this initial period.

The primary outcome of interest in this study was retention to the second year; a student was considered retained if they were enrolled in any course at the institution 12 to 18 months after their first course start date, or if they had earned a credential or graduated at any time between their first course start date and 18 months later. Progression during a student's first eight to nine months was measured by credit ratio which was operationalized as the number of credits earned with a grade of C or better divided by the number of credits attempted during the time period. Credits attempted during this time period were also recorded for each student.

Additional variables that could account for differences in retention or progression, such as student demographic and other academic factors, were explored and used as control variables providing greater confidence that the results concerning retention were related to delivery mode rather than other variables.

Methods of Analyses

Exploratory analysis was conducted comparing retention rates for three different groups of students based on their course-taking behaviors (delivery modes) in the first (approximately) eight to nine months at the institution. The initial exploratory analyses also compared differences in retention among the three delivery modes by Pell recipients, student age at entry, and gender.

Credit ratios and credits attempted for students taking only onground, only online, or blending their courses were additionally recorded and descriptively compared.

Because this was not a controlled experiment, there was concern that differences in retention rates among students in the different delivery mode groups could be due to inherent differences among those students, rather than an effect of their chosen delivery mode. To address this issue, variables that did not directly measure student success but had significant associations with retention at each institution were controlled for in a logistic regression model. Such variables affecting retention were identified at each institution individually and controlled for.

The effect of delivery mode was then added to the model to estimate the true relationship between course delivery mode and student retention. Variables that directly measure academic performance, such as credit completion, were not included as controls since there was no way of determining whether or not higher credit completion was a result of delivery mode or if the delivery mode chosen by a student was a result of the student being more academically gifted or motivated. However, course load (credits attempted), rather than course completion was controlled for in this study.

The following variables were considered as potential confounding factors:

- Military status at entry
- Pell status
- GED status
- Gender
- Race
- Degree sought at entry
- Student type at entry (as indicated by institution)
- Transfer college type
- Veteran at entry
- Credits attempted
- Developmental education ratio
- Student age at entry
- Median income associated with a student's home zip code
- High School GPA

Any of the above factors found to be significantly associated with retention at each individual institution were included as control variables in that institution's model. Two-way interactions between the above factors were also considered when their inclusion improved overall model fit. After appropriate control variables had been determined, the effect of delivery mode was added and measured as an odds ratio.

After calculating the average 12-18-month retention rates, average credit ratios, and credits attempted for students in each delivery group (blended, fully onground, and fully online) at each institution, averages for these three variables were aggregated across institutions in three categories: primarily onground community colleges, primarily onground four-year universities, and primarily online institutions. In each grouping of institutions, the averages for retention rates, credit ratio, and credits attempted give equal weight to each institution in the group, regardless of enrollment; that is, aggregated data was averaged with institution as the unit of analysis in an attempt to fairly account for the diversity represented by them.

Results

In the sections which follow results are given for each set of institutions—primarily onground community colleges, primarily ongound four-year universities, and primarily online institutions—relative to the research questions posed.

Primarily Onground Community Colleges

In this section, results are presented for students enrolled in five primarily onground community colleges and research questions related to them are explored. The community colleges in this group were public institutions located throughout the United States—in Florida, Ohio, Texas, Washington state, and Hawaii—and ranged in number of students enrolled for the period studied from almost 23,000 to just over 93,000. They also differed in the percentages of students who had Pell grants (41%

to 63%). Females outnumbered males at all the primarily onground community colleges studied. Traditionally aged students (< 26 years old) accounted for nearly three times as many community college students as older students

Do community college students who enroll in online courses have poorer course completion rates and are they retained at lower rates than community college students who take all their courses onground?

Table 1 shows the average retention percentages, credit ratios, and credits attempted for students enrolled in the five PAR community colleges considered in the analysis.

Table 1:*Average retention, credit ratios, and credits attempted for community college students by delivery mode*

	N	12-18 months	credit ratio	credits attempted
blended	91,622	58%	0.67	19.2
fully onground	112,269	51%	0.64	16.8
fully online	9,165	30%	0.64	10.2

The data show that across institutions, community college students taking some of their courses online and some onground had a retention rate of 58%, students taking all of their course onground had a retention rate of 51%, while only 30% of community college students taking all of their courses online were retained in the year following their first enrollment.

After controlling for possible confounding variables, logistic regression found that both the blended group and the fully onground group were slightly more likely to be retained than students who were fully online at all but one PAR community college. Odds ratios indicated that students blending their courses had 1.2 to 1.6 times greater odds of being retained than fully online students, and that fully onground students had 1.3 to 1.6 times greater odds of being retained than fully online students (Table 2).

There were no statistical differences between students in the blended group and fully onground students for four of the five institutions, indicating that observed differences in retention can be explained by the control variables rather than differences in delivery mode. For the one institution that did show a significant difference in odds between the blended group and the fully onground group, fully onground students had just 1.1 times the odds of being retained. The odds ratios for each institution are listed in the chart below. The retention odds of the delivery group listed first in the pairs were used as the numerators for the odds ratios.

Table 2: *Odds ratios comparing the odds of retention to a second year of students in differing delivery mode groups*

	Institution 1	Institution 2	Institution 3	Institution 4	Institution 5
fully onground vs. blended	1.0	1.0	1.0	1.1	1.0
blended vs. fully online	1.5	1.6	1.5	1.2	1.0
fully onground vs. fully online	1.6	1.5	1.6	1.3	1.0

The results reveal that students blending their courses attempted more credits on average (19.2) than either students taking solely onground (16.8) or solely online (10.2) courses.
Interestingly, there was very little difference in the average credit ratios (credits of C or better/credits attempted) among groups.

The results suggest that while taking all courses online had a mild negative impact on PAR community college students' retention, taking some online courses (blended) did not.
These results are distinctly different from the CCRC results (Jaggers & Xu, 2010; Xu & Jaggers, 2011) which suggested that taking any online courses hurt community college students' retention and progression. It may be that separating students taking only online courses from those taking some online courses (as done in this study) is what made the difference. The retention rates for community college students taking only online courses in the current study are quite low. It is also important to note, however, that less than 5% of community college students at the PAR institutions studied took all their courses online.

Moreover, credit ratios (which in some sense quantify pass rates) are remarkably similar across delivery modes in this study, indicating that community college students taking some or all of their courses online were as likely to complete and pass their courses as students taking all of their courses onground. Indeed, students blending online and onground courses were slightly more likely to complete their courses with passing grades than either students taking all their courses onground or students taking their courses only online. These findings seemingly contradict not only those of the CCRC (Jaggers & Xu, 2010; Xu & Jaggers, 2011), but those from the California community colleges (Hart, Friedman, & Hall, 2015; Johns & Cuellar Majia, 2014). The differences may be definitional. The credit ratio used in this research is just that: a ratio; it needs to be multiplied by credits attempted to get the number of courses completed with a passing grade. As the fully online students in the community college population studied in this research attempted far fewer credits, they would have obtained far fewer as well. Thus the small number of fully online students could bring down the "ever online" averages if one did not separate students who take all online courses from students who take some online courses, as was done in this study. In any case, the results clearly deserve further investigation.

Does delivery mode differentially affect particular groups of community college students?

To test whether delivery mode particularly affected different sorts of students, average 12-18-month retention rates were aggregated by delivery modes for students receiving or not receiving Pell grants, for female and male students, and for student age at entry (under age 26 or 26 years and older) across PAR primarily onground community colleges.

Table 3: *Community college student retention by Pell status*

	NO PELL		PELL	
	N	12-18 mo. retention	N	12-18 mo. retention
blended	47,117	53%	37,211	62%
fully onground	54,758	47%	39,843	57%
fully online	5,674	22%	3,426	42%

Table 3 shows the average retention percentages by Pell status for PAR students enrolled in community colleges. The data shows that students with Pell grants were retained at higher rates than those who did not have them, which may indicate the importance of financial support for such students. Moreover, the overall patterns of retention percentages within delivery mode groupings were somewhat different for students receiving and those not receiving Pell grants.
The difference between students taking only onground classes and those taking only online classes was considerably less for students with Pell grants than it was for students without Pell grants. The difference between students taking only onground classes and students taking some onground and some online classes, however, was essentially the same.

Table 4: *Community college student retention by gender*

	FEMALE		MALE	
	N	12-18 mo. retention	N	12-18 mo. retention
blended	54,179	60%	37,810	56%
fully onground	57,439	55%	54,488	49%
fully online	6,167	34%	2,987	25%

Table 4 shows the average retention rates for female and male students enrolled in PAR community colleges. They show that more women were enrolled and that they were retained at slightly higher rates than men. It is also interesting to note that women were considerably more likely than men to take any online classes. Just over half (51%) of the community college women in this study took at least one online course as compared with only 43% of the men, perhaps because more men are enrolled in technical classes only offered onground. The patterns of retention by delivery mode, however, were similar, indicating that delivery mode effects were not affected by gender.

Table 5: *Community college student retention by age*

	YOUNGER (< 26 years)		OLDER (26+ years)	
	N	12-18 mo. retention	N	12-18 mo. retention
blended	66,239	58%	25,383	57%
fully onground	86,512	53%	25,757	47%
fully online	4,819	26%	4,346	33%

Table 5 compares retention for each of the delivery modes between students 25 years of age and younger and students older than 25. A greater percentage of older community college students took only online courses, perhaps indicating their greater need for them, and older students taking only online courses were retained at much higher rates than younger students taking only online courses. In contrast, a greater percentage of younger community college students took only onground courses, and they were retained at much higher rates than older students taking only onground courses.

The results of these comparisons, then, suggest that while delivery mode did not differentially affect students grouped by gender, students grouped by Pell status were somewhat affected. Younger vs. older community college students were clearly differentially affected by delivery mode. The data suggest that taking only online classes is more harmful to younger students, and those without Pell grants in terms of observed retention rates. Potential differential effects of taking classes online should be explored for other student populations, especially at- risk groups such as ethnic and racial minorities, or students who are the first in their families to attend college.

Primarily Onground Four-Year Colleges

In this section, results for students enrolled in five, primarily onground, four-year colleges are presented and research questions related to them investigated. The analyses were initiated to see whether results for students taking online courses at community colleges also apply to students taking online courses at four-year universities.

The institutions represented in these analyses are public universities located throughout the United States—in Florida, Illinois, North Dakota, Arizona, and Hawaii—and ranged in number of students enrolled for the period studied from just over 3,000 to almost 60,000. Fewer students had Pell grants at these institutions than at the community colleges studied (18% to 42%) and females outnumbered males at all but one of the primarily onground four-year colleges studied. There were also slightly more traditionally aged students enrolled in the four-year universities studied.

Do students enrolled in four-year colleges who take online courses have poorer course completion rates and are they retained at lower rates than four-year college students who take all their courses onground?

Table 6 shows the average retention rates, credit ratios, credits attempted for PAR students enrolled in four-year colleges offering primarily onground programs. There were fewer university students than community college students in the PAR database, and there were slightly more students in this population blending their classes than students taking only onground classes. Students taking only online classes accounted for only a small percentage (6.5%) of these students, but it is interesting to note that combining students blending their classes with students taking only online classes reveals that the majority of students in this population were taking at least some online classes (ever online), which was not the case among the primarily onground community college population.

Table 6: *Retention rates, credit ratios, & credits attempted for primarily onground 4-year college students by delivery mode*

	N	12-18 mo. retention	credit ratio	credits attempted
blended	54,867	79%	0.80	24.4
fully onground	53,544	76%	0.81	25.2
fully online	7,625	60%	0.74	14.5

The data in Table 6 also show that students at primarily onground four-year universities were retained at higher percentages than primarily onground community college students in the PAR database. Moreover, although differences in retention percentages between delivery mode groupings were not as pronounced for the university population, the ranking patterns were the same. Students taking some of their courses online and some onground had a retention rate of 79%, students taking all of their course onground had a retention rate of 76%, and students taking all of their courses online had a retention rate of 60%.

After controlling for possible confounding variables in a logistic regression model, results showed that in the majority of cases no group was at higher risk than any other of not being retained at twelve to eighteen months. In one instance, students in the blended and fully onground groups had greater odds of being retained than fully online students with odds ratios of 1.6 and 1.8 respectfully. Significant differences were observed between students taking a blend of courses and students taking all of their courses onground at two institutions, but the effects were split: at one institution the fully onground group had only .77 times the odds of being retained as the blended group, while the fully onground group had 1.1 times greater odds of being retained at the other institution. Overall, the results suggest that the extraneous factors controlled for at each institution accounted for most or all of the differences in retention, rather than the effect of delivery mode. There was little evidence that taking any courses online at a primarily onground four-year university placed a student at greater risk of not being retained, once other variables are accounted for.

Odds ratios resulting from modeling the effect of delivery mode on retention with additional variables controlled for are shown in Table 7.

Table 7: *Odds ratios comparing the odds of retention to a second year of students in differing delivery mode groups*

	Institution 1	Institution 2	Institution 3	Institution 4	Institution 5
fully onground vs blended	.77	1.0	1.0	1.1	1.0
blended vs fully online	na*	1.0	1.0	1.6	1.0
fully onground vs fully online	na*	1.0	1.0	1.8	1.0

*not enough fully online students at Institution 1 for valid comparison

Credit ratios were higher for students at four-year universities in this study than for students at community colleges, and although the average credit ratios for students blending their courses and students taking all their courses onground were remarkably similar (.80 and .81 respectively), the average credit ratio for students taking only online courses was considerably smaller (.74). Similarly, the number of credits attempted by students taking only onground and students taking onground and online courses were similar (25.2 and 24.4 respectively) but students taking only online courses at primarily onground four-year universities attempted far fewer credits (14.5). It is possible that students taking only online courses chose to do so because they had busier lives than students taking some or all of their courses onground. Perhaps these students attempted fewer courses and had a more difficult time passing them for the same reason. The finding clearly deserves further investigation.

Does delivery mode differentially affect particular groups of four-year university students?

As with the community college data, further analyses were made to see whether Pell status, gender or age differentially affected four-year students in the PAR database. Indications were that they did not.

Table 8: *Four-year college student retention by Pell status*

	NO PELL		PELL	
	N	12-18 mo. retention	N	12-18 mo. retention
blended	40,310	78%	15,728	79%
fully onground	42,705	76%	18,283	74%
fully online	4,627	60%	2,998	62%

Table 8 shows retention percentages for students who did and did not receive Pell grants. Proportionally fewer university students than community college students received Pell grants in this study. Moreover, the retention rates are quite similar for those receiving or not receiving Pell grants, indicating that financial support may not be as important for retention among this group as it is with community college students. The overall patterns of retention percentages by delivery mode, however, were much the same as those for four-year college students in general, which suggests that Pell status did not differentially affect delivery mode differences.

Table 9: *Four-year college student retention by gender*

	FEMALE		MALE	
	N	12-18 mo. retention	N	12-18 mo. retention
blended	33,043	79%	21,820	78%
fully onground	26,569	77%	26,975	75%
fully online	4,781	60%	2,027	59%

Table 9 shows the average retention percentages for female and male PAR students enrolled in four-year colleges broken out by delivery mode. It is interesting to note that the majority of men enrolled in four-year universities chose to take all their courses onground, whereas only 41% of the women in this population did so. Indeed, 7% of the four-year university women but only 4% of the men took all their courses online. This finding deserves further explanation.

As with the community college data, Table 9 shows that more women were enrolled than men and that they were slightly more likely to be retained, but the gender differences were much smaller for the university populations. The patterns of retention by delivery mode, however, remained the same for this breakout, indicating that gender did not differentially affect differences related to delivery modes.

Table 10: *Four-year college student retention by age*

	YOUNGER (< 26 years)		OLDER (26+ years)	
	N	12-18 mo. retention	N	12-18 mo. retention
blended	14,551	78%	2.862	75%
fully onground	31.854	75%	2,402	70%
fully online	1,539	57%	3,605	57%

Table 10 presents delivery mode breakouts compared between students who were 25 years of age or younger and students who were 26 years of age or older. Due to technical problems, the data used in the comparison had to be reduced by one institution, the largest primarily onground university and one with a majority of students blending their classes. Thus it must be kept in mind that these age comparisons differ from the other comparisons in this category in that the percentage of students blending courses in the data used for the comparison was 15 percentage points fewer than that used for the other analyses in this category.

Nonetheless, the percentage of fully online students in the older group was ten times greater than the percentage of fully online students in the younger group, while the percentage of younger students taking all their courses onground was nearly three times as great as those taken by older students. This dichotomy suggests that the access to higher education provided by online courses has importance for older students. However, the patterns of retention in this category were quite similar across age groupings, indicating that delivery mode did not differentially affect students of differing ages.

The results of these comparisons, then, suggest that delivery mode did not differentially affect students receiving vs. students not receiving Pell grants, female vs. male students, or younger vs. older students at the primarily onground four-year universities involved in this study. Potential differential effects of taking classes online, however, should be explored for other student populations, especially at-risk groups such as ethnic and racial minorities, or students who are the first in their families to attend college.

Primarily Online Institutions

Finally, analyses similar to those undertaken for primarily onground community colleges and four-year universities were undertaken for four institutions whose primary focus was on online courses. These institutions were especially diverse. They included a public community college in the southwest, a public four-year university on the east coast, and two for-profit universities. The "primarily online" category was so constituted, in contrast to the more common practice of creating a unique "for-profit" category, because within the PAR community, we have found that institutions focusing on online programs are more similar to each other and less similar to primarily onground institutions.

The institutions representing the primarily online group had onground campuses located throughout the United States, and in one case, overseas. They had the largest number of students enrolled during the period studied, ranging from over 68,000 to almost 100,000. Percentages of students receiving Pell grants at these institutions ranged widely (20% to 80%) and females outnumbered males at all but one of the primarily online institutions studied.

The vast majority of the students in the primarily online category were, unsurprisingly, fully online students. However, three of the four institutions in this group had a sufficient number of fully onground students to make it possible to investigate all three levels of delivery mode, making it possible to ask similar questions of this category of institutions. Part of the rationale for doing so was to explore whether primarily online institutions might have better retention rates for students taking online courses than primarily onground institutions. They did not. The patterns of retention across delivery modes, however, were somewhat different.

Are there any differences in completion and/or retention rates associated with differing delivery modes at primarily online institutions?

Table 11: *Retention and credit ratios for students enrolled in primarily online institutions by delivery mode.*

	N	12-18 mo. retention	credit ratio	credits attempted
blended	29,577	45%	0.66	15.7
fully onground	48,067	29%	0.76	13.4
fully online	252,522	31%	0.52	9.7

Table 11 shows the 12-18-month retention rates, credit ratios and credits attempted for students enrolled in primarily online institutions broken out by delivery mode. The retention rates are clearly lower for this category of institutions except among students taking all their courses online, for whom the retention rates are similar to those of community college students taking all their courses online. Table 9 also shows that students taking only online courses in primarily online institutions were slightly more likely to be retained than students taking only onground courses, a deviation from the patterns across delivery modes for primarily onground institutions. However, as with primarily onground institutions, students taking some courses online and some courses onground were retained at higher rates than students taking only onground or only online courses.

After controlling for possible confounding variables in a logistic regression model, results revealed moderate differences between students taking a blend of courses and students taking their courses exclusively online or exclusively onground (Table 12). Students in the blended group at primarily online institutions had between 1.2 and 1.8 times greater odds of being retained than students taking all their courses onground (although there were no significant differences at one such institution and not enough only onground students to make a comparison at another). Students in the blended group also had between 1.2 and 1.4 times greater odds of being retained than students taking all their courses online (although there were no significant differences between the groups at one institution). Students taking only onground courses had greater odds of retention than students taking only online courses at one institution and slightly lower odds of retention than students taking only online courses at another.

Table 12: *Odds ratios comparing the odds of retention to a second year of students in differing delivery mode groups*

	Institution 1	Institution 2	Institution 3	Institution 4
blended vs fully onground	na*	1.8	1.2	1.0
blended vs fully online	1.3	1.4	1.2	1.0
fully onground vs fully online	na*	.77	1.0	1.2

*not enough fully onground students at Institution 1 for valid comparison

The results suggest that among the PAR primarily online institutions, students taking a blend of courses were more likely to be retained, but there is little evidence that there are significant differences in retention odds between fully onground and fully online students after accounting for extraneous variables. The results suggest that at primarily online institutions, students taking some courses online and some courses onground are more likely to be retained than students taking all their courses either online or onground. While it should be noted that only a small percentage (9%) of these students blended their courses, the finding clearly deserves further investigation.

The average credit ratio for students blending courses at primarily online institutions (.66) was similar to that for students blending courses at community colleges, while it was lower than that for students taking only online courses at community colleges, and for all delivery categories at four-year universities. The average credit ratio for students at primarily online institutions taking only onground courses was quite high (.76) which is interesting, but may reflect the particular nature of onground courses at online institutions. It may be that the onground courses offered include a significant number of orientation courses or success type courses. Clearly, however, the result should be further investigated.

The average credits attempted by students enrolled in primarily online institutions were a good bit lower across all categories than average credits attempted at either community colleges or four-year universities. This finding suggests that students enrolling at primarily online institutions may do so because of access issues, most likely time constraints. Time constraints might explain not only lower credits attempted but also lower retention and credit ratios as well. This notion too clearly deserves further investigation.

Does delivery mode differentially affect particular groups of college students at primarily online schools?

As with the primarily onground institutions, further analyses were made to see whether Pell status, gender or age differentially interacted with the ways in which delivery mode affected students at primarily online PAR institutions. Indications were that they did not, except in one somewhat questionable instance.

Table 13: *Student retention at primarily online institutions by Pell status*

	NO PELL		PELL	
	N	12-18 mo. retention	N	12-18 mo. retention
blended	20,933	40%	25,344	51%
fully onground	33,453	25%	14,614	61%
fully online	64,880	25%	141,102	39%

Table 13 compares students who received Pell grants with students who did not receive them for all three delivery modes. A higher percentage of students in the primarily online segment received Pell grants when compared to the primarily onground institutions, and, as with community college students, Pell recipients were retained at higher rates, indicating perhaps a greater need for financial assistance. With one exception, the patterns of retention for each delivery mode resembled that of overall retention by delivery mode for the primarily onground category. The one exception was a higher rate of retention for Pell recipients taking only onground classes than for students in any other delivery mode by Pell status category. Such students only represent about 4% of the total population and the difference seems to be driven by a single institution with a tiny onground population (and so primarily an artifact of the way averages were calculated). Thus, while it is clearly of interest to that institution, it probably does not represent a general trend.

Table 14: *Student retention at primarily online institutions by gender*

	FEMALE		MALE	
	N	12-18 mo. retention	N	12-18 mo. retention
blended	26,939	49%	18,744	43%
fully onground	21,488	30%	25,109	26%
fully online	163,741	34%	93,501	28%

Table 14 shows retention broken out by gender for students in all three delivery mode groupings—blended, fully onground, and fully online—at primarily online institutions. The data show that there were considerably more women than men enrolled at such institutions and that women were retained at slightly higher rates than men. The patterns of retention by delivery mode, however, mirror those for this category of institutions overall.

Table 15: *Primaily online student retention by age*

	YOUNGER (< 26 years)		OLDER (26+ years)	
	N	12-18 mo. retention	N	12-18 mo. retention
blended	20,782	41%	25.487	47%
fully onground	48,167	28%	74.158	27%
fully online	95,778	27%	156.683	34%

Table 15 explores possible differential effects of age on the retention of students in differing delivery mode categories. There were many more older students in this category than younger, which is quite different from students in primarily onground institutions. Moreover, older students taking only online courses in this category were retained at a considerably higher rate than younger ones taking only online courses, another indication of the importance of this delivery mode for older students.

The results of the comparisons in this category suggest that delivery mode did not differentially affect students receiving vs. students not receiving Pell grants or male students vs. female students at the primarily online institutions involved in this study. However, age differences did seem to do so. Potential differential effects of taking classes online should be explored for other student populations, especially at-risk groups such as ethnic and racial minorities, or students who are the first in their families to attend college.

Discussion

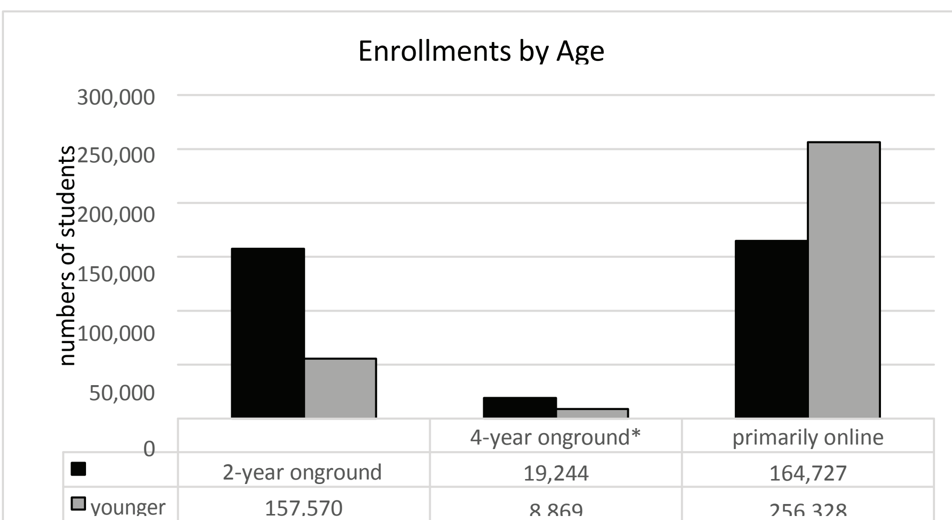
To a large extent, online learning was developed to provide access to higher education for underserved populations, and access for underserved populations remains a measure of quality in online programs (Online Learning Consortium, 2015). Some of the findings in this study provide evidence that online courses and programs continue to offer such access.

Table 16: Enrollments by institution type and delivery mode

	onground CC	4-year onground	primarily online	totals
blended	91,622	54,867	29,577	176,066
fully onground	112,269	53,544	48,067	213,880
fully online	9,165	7,625	252,522	269,312
totals	213,056	116,036	330,166	659,258

For example, there were many more students enrolled in the primarily online institutions investigated in this study (Table 16), than in either the community colleges (more than 1.5 times as many) or the four-year universities (2.8 times as many). Although the institutions in this study are not a representative sample of institutions across the nation, the raw number of students in the primarily online segment demonstrates the current demand for online courses and programs. It is also important to note that none of these institutions were primarily online 20 years ago, that they have all grown phenomenally as a result of online offerings, and that such growth has only been possible because of the scalability of online courses. One could also argue that these institutions have grown because there is a demand for the access online courses provide to students with busy lives.

Indeed, a higher percentage of older students were enrolled at primarily online institutions (Figure 1). There were more than 1½ times as many older students as younger ones at primarily online institutions, while there were nearly three times as many younger students at primarily onground community colleges. It thus seems that online courses offer particular access to higher education for older students in a way similar to the access given to younger students by community colleges.



*4-year onground institutions in this comparison include 4/5 of the schools in other comparisons

Figure 1: A comparison of enrollments by age across differing categories of institutions

In addition, considerably more than half of all students at primarily online institutions received Pell grants, whereas at primarily onground institutions more than half of all students did not receive Pell grants (Figure 2).

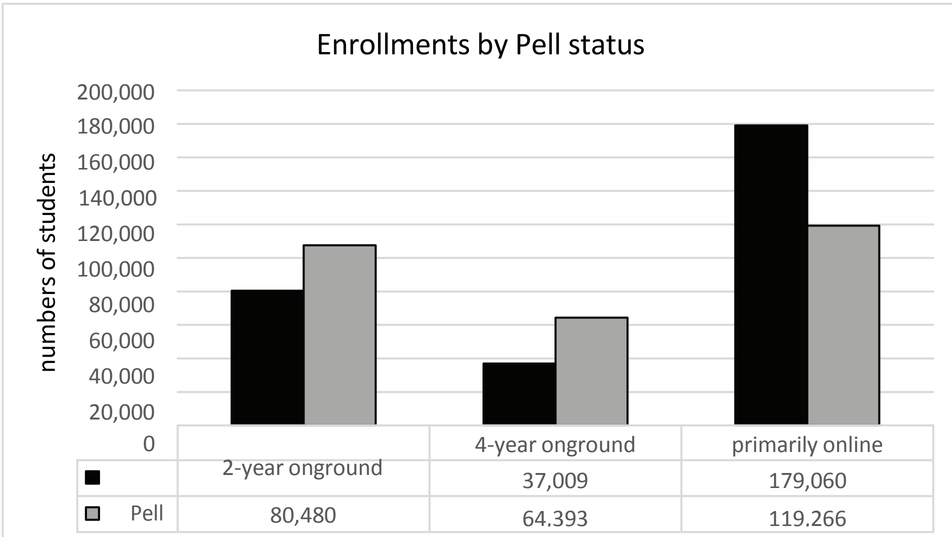
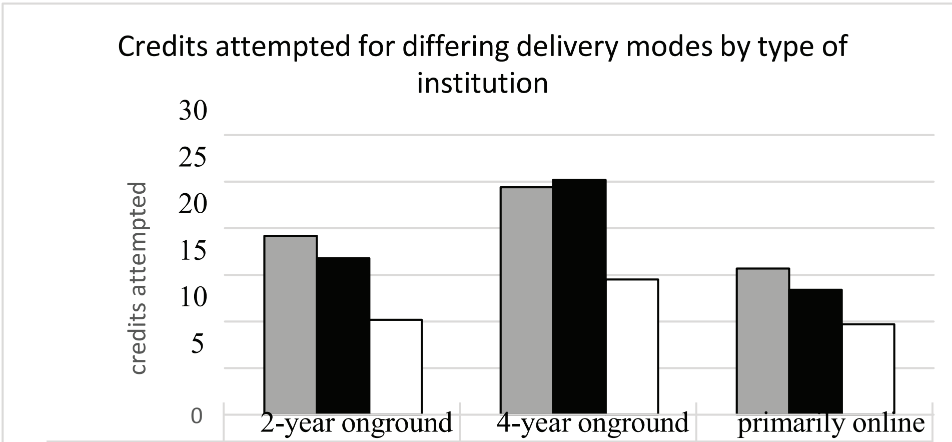


Figure 2: A comparison of enrollments by Pell status across differing categories of institutions The growing number of students who informally blend online and onground courses

(Bloemer & Swan, 2014) also suggests the importance of online offerings for those students we continue to call "alternative." Almost half the students at primarily onground community colleges and over half the students at primarily onground universities took at least some courses online in their first eight months. In fact, "traditional" students, who attend college full time right after high school and live on campus, only amount to 15% of those attending college in this country, while alternative students -- "working parents, veterans and military personnel, caregivers and others" -- make up 85% of that population (UPCEA, 2014).

Another piece of evidence which points to online courses being a way for nontraditional students to access higher education is a comparison of the number of courses students are taking at different categories of institutions in differing delivery modes. While credit ratios (Figure 3) are not all that different across institutional types, the numbers of credits attempted really are quite different (Figure 4). Across institutional types, the average number of credits attempted for both students taking only onground courses and students blending online and onground courses was 19.8. The average number of credits attempted for students taking only online courses was 11.5. On the other hand, it is interesting how remarkably similar the patterns of credits attempted look for community college students and students attending primarily online institutions in this study. One could argue that these findings again point to the notion that community colleges have been the traditional way for nontraditional students to access higher education, and that primarily online schools are emerging as another path to higher education for nontraditional students. At the very least, both are open access institutions.

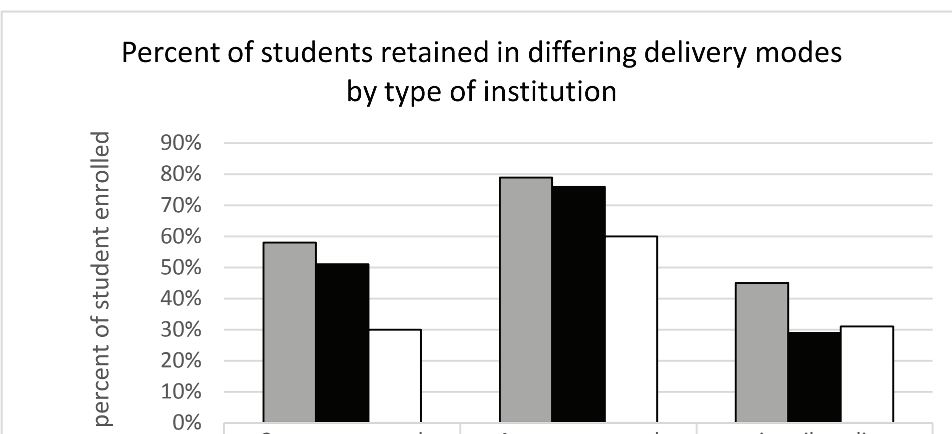
Figure 3: A comparison of credit ratios by delivery modes across differing categories of institutions



	2-year onground	4-year onground	primarily online
blended	19.2	24.4	15.7
Fully onground	16.8	25.2	13.4
fully online	10.2	14.5	9.7

Figure 4: A comparison of credit attempted by delivery modes across differing categories of institutions

If a primary intention of online learning is to provide educational opportunities for students who are not able to attend courses in person due to time and/or location restraints, it is important to keep that intention in mind. The results of this study suggest that when other student characteristics are accounted for, participation in online and onground courses did not have a large impact on odds of retention when compared to participation in only onground courses. In fact, in most instances students who blended online and onground courses were more likely to be retained than students taking only onground courses.



	2-year onground	4-year onground	primarily online
blended	58%	79%	45%
fully onground	51%	76%	29%
fully online	30%	60%	31%

Figure 5: Comparisons of retention rates by delivery modes and institution types

This is an important finding in that it suggests that the widely circulated notion that taking any online classes hurts community college students in particular, and other post-secondary students by implication, is not necessarily the case. For community college students in this study, taking some online classes led to slightly higher retention and progression rates than taking only onground classes (Figure 5). Moreover, the same was true of students enrolled in primarily onground four-year universities and primarily online institutions, and across students differentiated by Pell status and gender.

Why blending courses is associated with higher retention rates surely deserves further investigation. Perhaps blending courses gives students greater flexibility than taking courses only onground, and so makes it easier for them to stay enrolled when things in their lives might interfere with onground schedules. On the other hand, perhaps blending courses gives students a feeling of attachment to a real place, a brick and mortar institution, and that is more binding than the attachment to a virtual institution associated with taking courses only online. Understanding what makes blending courses a more effective delivery mode might help us better support all students.

Findings from this study also indicated that taking all courses online was at worst associated with moderately lower odds of retention (the largest odds ratio being 1.8 when compared against fully onground students) and at best associated with the same odds of retention as students taking all courses onground or students blending their courses online. It is important to note in this regard, that, if online courses are the only way some students can access higher education, deployed military or single mothers for example, then they are really not bad at all.
Perhaps finding ways of adding some onground experiences to primarily online programs would further reduce any differences in student retention.

Indeed, differences observed in retention rates between the three different delivery modes were largely explained by extraneous factors, rather than the delivery mode itself. One such extraneous factor was total credits attempted, which may be associated with the available time a student has to participate in courses. Lower credit attempts were negatively associated with retention at the participating institutions, and fully online students also attempted fewer credits than the students in the blended and fully onground groups at the primarily onground institutions. In this way, it appears online courses likely serve otherwise at risk students, who may among other things have time constraints that make attending onground courses difficult. Further points of study may be to investigate what factors place online students at risk that may not be risk factors for fully onground students, and to investigate whether specific segments of the student population are better suited for online or onground coursework.

Another area for investigation is institutional differences in the way students experience online course-taking and the degree to which certain institutional policies and practices related to the timing of online courses impact course completion and retention. Even within an institutional grouping with consistent trends, such as with traditional community colleges, odds ratios related to retention vary. Within our PAR member institutions, policies related to timing of online course-taking also differ. For example, at one of the traditional four-year schools, students are not allowed to enroll in online courses until their second semester.

Both primarily online and traditional onground schools often use college readiness instruments that measure technology aptitude to determine if a skills gap needs to be filled before embarking on online course-taking. Most institutions that offer online courses have online orientations – some mandatory, some optional- to help students learn more about the institution's particular approach to online courses prior to taking a credit bearing course. All these factors
and other interventions designed to enhance online success merit additional research to determine their contribution and influence on course outcomes and retention.

Limitations

Data for this study was limited to the fourteen PAR Framework institutional partners who agreed to participate, and as such represents a sample of convenience rather than a stratified national or regional sample of post-secondary institutions. Findings, therefore, may or may not be representative of undergraduate students nationally. The participating institutions, however, were quite diverse both in student bodies and institutional focus, and they were located throughout the United States. Although an extensive number of potential confounding factors were considered and controlled for, it is also possible that unmeasured variables could have influenced the results as well. For instance, information on whether or not a student has a full time job and/or is providing for a family could impact the results as such students would likely have less time to spend on coursework and may be more inclined to take courses online as a result. It could be that some of the negative relationships observed between online course-taking and retention may actually reflect that many online students have less time to dedicate to their studies, and more information on a student's work and family life could help control for those potential differences. This research is also limited to retention to a second year. A second and more extensive study could include retention beyond the first year and the examination of graduation rates among students in the three different delivery modes.

Conclusions

The research reported in this paper found that, contrary to what has been widely reported in the press taking some online courses did not result in lower retention or course completion rates for students enrolled in primarily onground community colleges participating in the PAR Framework. Moreover, although retention rates were lower for such students taking only online courses than for similar students taking only onground or blending their courses, the odds ratios for these differences were small (Chen, Cohen, & Chen, 2010).

One important explanation for differences between these finding and other reports (Hart, Friedman, & Hill, 2015; Jaggers & Xu, 2010; Johnson & Cuellar Majia, 2014; Xu & Jaggers, 2011) is the way the data was grouped. In previous reports, students who took some online courses were grouped with students who took only online courses as "ever online". In the current study, these groups were separated with illuminating results – students who took some (but not all) courses generally were retained at higher rates than students who took all their courses onground; while students who took all their courses online had lower rates of retention. The findings suggest that taking some online courses is definitely not harmful and indeed may be beneficial. They perhaps provide some explanation for Shea and Bidjerano's (2014) findings associating taking some online classes with more likelihood of receiving a credential. The findings and the methodology surely deserve further investigation.

Essentially no differences in retention between delivery mode groups were found for students enrolled in primarily onground four-year universities participating in the PAR Framework, while at participating primarily online institutions, students blending their courses had slightly better odds of being retained than students taking exclusively onground or exclusively online courses. No differences between the latter groups were found at these institutions.

This study extends explorations of retention and the taking of online classes to the four-year and primarily online populations and finds no significant problems associated with taking online courses. The latter grouping, "primarily online," is another categorization unique to this study which may be significant. The study found that primarily online institutions, whether they be community colleges, four-year public institutions, or for-profit colleges, are more like each other than primarily onground institutions. In particular, they serve a much larger alternative, and probably "at-risk" population. This notion should be investigated further.

Moreover, patterns of retention were the same regardless of gender across institutional categories and almost the same regardless of Pell status (there was a greater difference between students with a Pell grant and students without one among fully online students at primarily onground community colleges, than observed in the other two delivery modes). Age did seem to be a factor differentially affecting delivery mode at primarily onground community colleges and primarily online institutions. Older community college students taking only online courses were retained at higher rates than younger students taking only online courses. Similarly, older students taking only online courses or taking some online and some onground courses at primarily online institutions were retained at higher rates than younger students in the same categories. No differential effects of delivery modes were found for any groupings (female vs. male, Pell vs. no Pell, or younger vs. older students) among students enrolled in primarily onground four-year universities in this study.

The results thus suggest that online courses may provide access to higher education for people who have not traditionally had such access. Future investigations should examine the taking of online classes among traditionally underrepresented populations. The results indeed indicate that taking online courses is not harmful for most students, and may in fact be beneficial when online courses are blended with onground courses. The issue clearly deserves further investigation.

References

Allen, I. E., & Seaman, J. (February. 2015). *Grade Level: Tracking Online Education in the United States.* Babson Survey Research Group.

Bernard, R., Abrami, P., Borokhovski, E., Wade, C., Tamim, R., Surkes, M., et al. (2009). A meta-analysis of three types of interaction treatments in distance education. *Review of Educational Research, 79*(3), 1243–1289

Bloemer, W. & Swan, K. (2014). Investigating informal blending at the University of Illinois Springfield. In A. G. Picciano, C. D. Dziuban, & C. R. Graham (Eds), *Blended Learning Research Perspectives, Volume 2.* New York: Routledge, 52-70.

Chen, H., Cohen, P., & Chen, S. (2010). How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in Statistics – Simulation and Computation, 39*(4), 860-864.

Hart, C. M. D., Friedmann, E., & Hill, M. (April, 2015). Online course-taking and student outcomes in California community colleges. Paper presented at the 2015 annual meeting of the American Educational Research Association, Chicago.

Jaggers, S. S., & Xu, D. (September, 2010). *Online learning in the Virginia Community College system*. New York: Community College Research Center, Teachers College, Columbia University.

Johnson, H., Cuellar Mejia, M., & Cook, K. (2015). *Successful Online Courses in California's Community Colleges*. San Francisco, CA: Public Policy Institute of California.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). Evaluation of evidence- based practices in online learning: A meta-analysis and review of online learning studies. U.S. Department of Education.

Moore, J. C., & Fetzner, M. J. (2009). The road to retention: A closer look at institutions that achieve high course completion rates. *Journal of Asynchronous Learning Networks, 13*(3), 3-22.

Online Learning Consortium. (2015). Five pillars of quality online education. Retrieved from: http://onlinelearningconsortium.org/about/quality-framework-five-pillars/

Shea, P. & Bidjerano, T. (2014). Does online learning impede degree completion? A national study of community college students. *Computers & Education, 75,* 103–111

UPCEA (University Professional and Continuing Education Association). (October, 2014). Federal Policy Brief. Retrieved from: http://www.upcea.edu/Files/UPCEAPolicyBrief.pdf

Xu, D., & Jaggers, S. S. (March, 2011). *Online and hybrid course enrollment and performance in Washington State community and technical colleges*. New York: Community College Research Center, Teachers College, Columbia University

Assessment of Learning in Digital Interactive Social Networks: A Learning Analytics Approach

Mark Wilson, Perman Gochyyev
University of California, Berkeley

Kathleen Scalise
University of Oregon

Abstract

This paper summarizes initial field-test results from data analytics used in the work of the Assessment and Teaching of 21st Century Skills (ATC21S) project, on the "ICT Literacy — Learning in digital networks" learning progression. This project, sponsored by Cisco, Intel and Microsoft, aims to help educators around the world enable students with the skills to succeed in future career and college goals. The paper begins with describing some expansions to a common definition of learning analytics, then includes a review of the literature on ICT literacy, including the specific development that led to the ATC21S effort. This is followed by a description of the development of a "learning progression" for this project, as well as the logic behind the instrument construction and data analytics, along with examples of each. Data were collected in a demonstration digital environment in four countries: Australia, Finland, Singapore and the U.S. The results indicate that the new constructs developed by the project, and the novel item forms and analytics that were employed, are indeed capable of being employed in a large-scale digital environment. The paper concludes with a discussion of the next steps for this effort.

Acknowledgements: We thank the ATC21S project and its funders for their support for the work reported in this report. We also acknowledge the expertise and creative input of the ATC21S Expert Panel in ICT Literacy: John Ainley (Chair), Julian Fraillon, Peter Pirolli, Jean-Paul Reeff, Kathleen Scalise, and Mark Wilson. Of course, the views and opinions expressed in this paper are those of the authors alone.

Introduction

The view of learning analytics that this paper is based on starts with the observation that the current practices of schooling are somewhat outmoded in the new global working environment. Today, people work both individually and in groups to share complementary skills and accomplish shared goals—this practice contrasts to schools and assessments where students take tests individually.

Knowledge is applied across disciplinary boundaries in the process of solving real-world problems, but in schools, the school subjects are divided by disciplinary boundaries. Furthermore, at work, problem solving is often complex and ill-structured, in contrast to the simplified problem-types featured in much of school education, and especially school standardized testing. Finally, in the work setting, people have access to enormous resources of information and technological tools, where the challenge is to strategically craft a solution process, which contrasts strongly with the traditional "closed book" analytics of what learners know and can do (CIM, 2008).

As a response to such transitions, the Assessment and Teaching of Twenty-First Century Skills project (ATC21S) was launched in 2009. A goal was to employ new analytical approaches to the assessment of learning such that the challenges above could be better addressed. We will utilize an example of work from this project to illustrate the points we are making in this paper.

A commonly used definition of learning analytics that we will draw on here was proposed by the first international Conference on Learning Analytics and Knowledge (LAK 2011) and adopted by the Society for Learning Analytics Research (Society for Learning Analytics Research, 2011):

> "Learning analytics is the 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."

While this definition is helpful, two additional aspects are important to consider: the interpretation of results, and the choice of appropriate data types and algorithms.

The paper begins with a description of some expansions to a common definition of learning analytics, then includes a review of the literature on ICT literacy, including the specific development that led to the ATC21S effort. This is followed by a description of the development of a "learning progression" for this project, as well as the logic behind the instrument construction and data analytics, along with examples of each. The paper goes on to describe the data that was collected, and reports on the learning analytics approach that was used. The paper concludes with a discussion of the next steps for this effort.

Bridging the Gap between Evidence and Interpretation: Some Background

First, in this paper we make the point that, for learning analytics, the meaningful interpretation of the data analysis is critical to consider, not simply reporting the results (Wilson, 2005; Wilson et al., 2010; Wilson, Scalise, & Gochyyev, 2014). Interpretation is not directly included in the LAK/SoLAR definition of "collection, analysis and reporting." This weakness in the definition can lead to the assumption that once results are composed and reported, their meaning for learners and learning outcomes is self-evident.

However, that said, learning analytics as described in the LAK/SoLAR definition constitute a type of educational assessment. As such, meaningful interpretation means having an evidentiary framework designed to connect results clearly and on an empirical basis back to the goals and objectives of the analysis in order to make clear evidentiary claims about the learner (Mislevy, Almond, & Lukas, 2003; Wilson, 2005). It also means being able to understand the uncertainty or range and degree of error likely to be present in the results.

Some groups have begun to establish standards of practice in learning analytics for such 21st century complex data analysis methodologies (Sclater, 2014; Wilson et al., 2012). In this paper, we will present an example that helps establish the coherent evidentiary argument for the learning analytics involved through a framework called a "learning progression." This framework connects the results to (a) the data and the learning analytics? questions being asked, and (b) to the techniques for the analytics employed. Other researchers have begun to describe the need for such frameworks when learning analytics goes beyond data analysis alone and is to be used for predictive analytics, actionable intelligence, and decision-making (van Barneveld, Arnold, & Campbell, 2012). In learning analytics, the need to establish a coherent evidentiary argument to support claims about learners can be approached either *a priori* (in advance of the analysis) or *a posteriori* (following the analysis). The *a priori* approach is essentially a theoretical approach, based on a strong theory or prior empirical information (or both), and thus might be considered a confirmatory learning analytic technique.

The *a posteriori* approach can be considered generative, or in other words, exploratory, and in many cases will need to be confirmed by a subsequent data collection and analysis. The exploratory approach is sometimes called by the name “data mining” (Papamitsiou & Economides, 2014). Exploratory approaches can be useful when the research goal is to learn more about the patterns in the data sets in a context where little is yet understood, or where new patterns may become evident that were not previously suspected.

The choice between an exploratory or confirmatory approach is an option, with the choice depending on how much prior theory and/or empirics are available. Put together, these exploratory and confirmatory stages can be seen as a cycle in the evidence chain, as shown in Figure 1. The figure depicts a simple example of a learning analytics interpretive cycle, where entry points can be either confirmatory—entering at the “theory or conceptualization” node—or exploratory—entering at “analysis and results” node—for extant data or at the “empirical data” node when observations will be designed and collected (see below for a discussion of extant and collected data).

No single entry point to the cycle is best among the choices in every situation. The choice can be determined by the intended *purposes* of interpretation, and the current *state of claims* that can be made in a given context. In any particular situation, one relevant question to ask is does the analysis begin with an interpretive framework *a priori,* as in the theory component of the cycle below, or is interpretation made *a posteriori*, as when even the initial interpretive framework is derived from data because little is yet known?

In either case, the same cycle is present but with different points of entry and a different flow to the interacting elements.

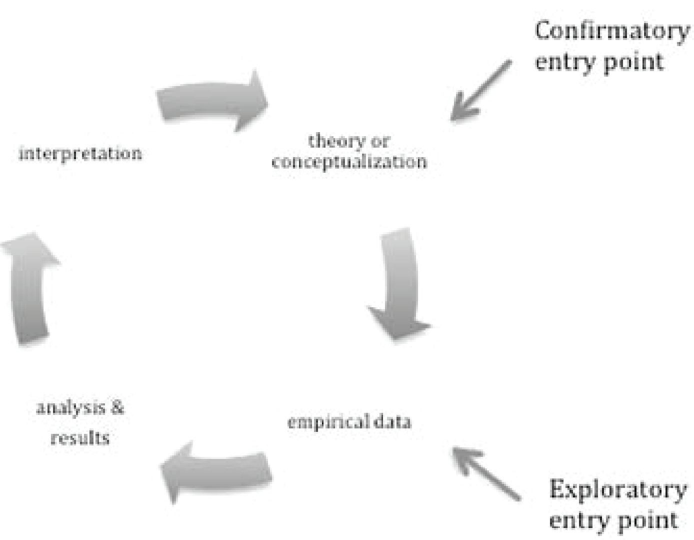


Figure 1. Choice of appropriate data types and algorithms

In terms of data types for which learning analytics by the LAK/SoLAR definition are likely to be useful, in most cases complex data should be involved. If not, other simpler techniques might be better employed (Ferguson, 2012). Complex data can take the form of large data sets (big data), multi-faceted data sets, or other elements in the data that encode more complex patterns or hard-to-measure constructs not readily identifiable without complex analytic techniques (Scalise, 2012; Wilson et al., 2010).

The unit of analysis itself can also generate data complexity, for instance, when data are collected for individuals but the individuals are collected into groups, programs, institutions, or other collective units, for which data are also collected.

Sometimes data sets can be pre-existing, or extant data sets, as described above. Examples of pre-existing data include downloads from Twitter feeds, click streams in user data, or other online collections generated as a result of processes with various purposes (Baker & Siemens, 2014). At other times, data sets are collected at least in part directly for the purpose of applying learning analytics to the results. Data collection can include, for instance, an adaptive recommender where ratings on prior experiences are solicited for the purposes of prediction of respondent interest in future experiences (Chedrawy & Abidi, 2006; Dagger, Wade, & Conlan, 2005), or evidentiary data collection for educational or professional development, to address personalized or grouped components to support the learner in educational assessment (Brady, Conlan, Wade, & Dagger, 2006; Kennedy & Draney, 2006).

An extension to the LAK/SoLAR definition we propose here is the specification that complex analytic techniques are needed to resolve the multifaceted or complex patterns. The same argument can be made as above for data sets. Complexity should be introduced in the analysis for a coherent evidentiary argument only when necessary. So the usual parsimonious definition should be applied when models or other algorithms are used to fit learner data and resolve patterns.

Finally, it would be helpful if the LAK/SoLAR definition made reference to algorithms, or characteristics of algorithms, that might be to useful to apply for aggregating and parsing of patterns, since this is an important consideration in the use of learning analytics (Papamitsiou & Economides, 2014). While it is important to keep the definition general to be inclusive of many useful algorithms that might arise, as a general class the approach typically needs to involve algorithms to automatically process the data, assuming the purposes of interpretation and the complexity of data require algorithmic approaches to the accumulation and parsing of patterns in the data. Algorithms can be statistical in nature, applied as inferential statistical tests or to yield inferential indices as part of the processing, which can help with assessing quality of results (Sclater, 2014). Numerous algorithms in the form of measurement models that take a statistical form for learning outcomes have been created and applied. These are well established in the psychometrics research literature and some of the advanced models as well as basic models can be appropriate to apply in learning analytics to complex 21st century skill settings (Wilson et al., 2012). Algorithms can also process patterns in more descriptive ways, yielding machine-readable results such as categorization or subsetting of respondents (Stanton, 2012). Note that since machine processing is required, however, the data sets at some point have to include machine-readable data. This may be text-based or graphical in nature, or in some other innovative format, depending on the processing requirements of the algorithm and platform, or the data sets may be numeric (Scalise & Gifford, 2006). The desired data characteristics may already be present for a given data set in any particular case, or may require preprocessing. This could include types of scoring, ordering, subsetting, or other types of aggregation. For this, reliable data collection, warehousing and prep can be a problem, so a variety of "clean-up" procedures may be needed. An important stage in learning analytics is reducing construct irrelevant variance including noise, user errors, or out-of-scope entry of data, which should be clarified and validated before conclusions can be drawn (Dringus, 2012).

In light of these sets of clarifications, we suggest a revision to the LAK/SoLAR definition, which we propose as "Learning analytics definition, LAK/SoLAR.v2":

"Learning analytics is the measurement, collection, analysis, *interpretation*, and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs, by means of a coherent evidentiary argument... Complexity should be introduced in the data and the analysis only when necessary to the development of the evidentiary argument." Complex data as described here can take the form of large data sets (big data), multi-faceted data sets, and/or other the data elements that encode patterns or hard-to-measure constructs not readily identifiable without advanced analytic techniques.

In the sections that follow, we introduce the ATC21S project, which is the context of our study, and review the literature on ICT Literacy, culminating in a section on the inception of the ATC21S ICT Literacy effort. We then give a very brief description of the BEAR Assessment System mentioned above, as that is used as the organizing principle for the next four sections: The construct map, the items design, the outcome space, and the measurement model. This is followed by a description of the key results from an empirical study of the assessments using a sample of students from four countries: Australia, Finland, Singapore and the United States. The paper finishes with a discussion of conclusions and next steps.

Overview of the ATC21S project

ATC21S was founded, in an unusual collaborative effort, by three information technology companies (Cisco, Intel, Microsoft—CIM). According to the company collaborative:

> The economy of leading countries is now based more on the manufacture and delivery of information products and services than on the manufacture of material goods. Even many aspects of the manufacturing of material goods are strongly dependent on innovative uses of technologies. The start of the 21st century also has witnessed significant social trends in which people access, use, and create information and knowledge very differently than they did in previous decades, again due in many ways to the ubiquitous availability of ICT. (CIM, 2008)

It is to be expected that this widespread change will significantly affect the personal and working lives of many people, and hence should have equally large effects on the school systems that educate people for their later lives and careers. These effects may include the nature of and even the names of the subjects that are taught in schools, how those new subjects (and the traditional subjects) are taught and learned, and how education is organized. Thus, the ATC21S project was initiated to develop new assessments in the area of 21st century skills, based on the idea that new assessments could lead the way to these new subjects. The project determined to develop sample assessments in a few specific domains of the 21st century skills, to serve as examples leading to further developments in learning analytics in those areas.

The project recruited a large group of experts in diverse fields to author a set of "white papers" that paved the way for further work. This process is documented in Griffin, McGaw & Care (2012), and two chapters from that volume are germane to the efforts delineated here: the "skills white paper" (Binkley, Erstad, Herman, Raizen, Ripley, Miller-Ricci & Rumble, 2012); and the "methodology white paper" (Wilson, Bejar, Scalise, Templin, Wiliam, & Torres-Irribarra, 2012). The skills white paper develops a framework for describing and implementing the 21st century skills and how they relate to traditional school subjects. This framework is labeled "KSAVE," for Knowledge, Skills and Attitudes, Values and Ethics. Two particular 21st century skills were chosen for inclusion in an ATC21S demonstration project—collaborative problem solving and ICT literacy. This second 21st century skill is the focus of this paper, and our particular approach to that construct is described in this paper. The "methodology white paper" describes a method to develop the new assessments (NRC, 2001). The approach chosen is called the BEAR Assessment System (BAS: NRC, 2001; Wilson, 2005; Wilson, 2009a; Wilson & Sloane, 2001), and it will also be briefly introduced here, though the reader is directed to the references just cited for a more complete account.

Brief Literature Review of the Domain

For this project, learning analytics were applied in an *a priori* approach, entering at the theory/conceptualization node of the learning analytics interpretive cycle discussed in the prior section, since considerable information about the domain is already available. Information about the domain is based on prior empirical results and theory briefly summarized in this section.

The idea of learning through digital networks and the use of digital media, in the context of schooling, has been based on the conceptualization of student use of information and communication technology, or "ICT" literacy. To wit, information technologies are seen as resources for creating, collecting, storing, and using knowledge as well as for communication and collaboration (Kozma, 2003).

The definitions of information and communication literacy range from simple—basic digital knowledge and skills—to complex, a wide range of tools and competencies that ATC21S describes as "tools for working" in a digital age (Binkley et al., 2012). These generalized tools include accessing and using technologies to evaluate and utilize information, analyzing media, creating information products, and they include understanding ethical and legal issues involved in being a participant in the knowledge economy.

Efforts to measure ICT practices in schools go back about three decades—this history has recently been reviewed by Wilson, Scalise, & Gochyyev, (2015) so we will not attempt to duplicate that history here. In that paper, the authors described how the concept of ICT literacy has changed a great deal during these years: From a conceptualization as a specific domain of knowledge about computers, to an understanding of it as a domain-general or transversal 21st century skill. They traced the conceptual changes in the idea of ICT literacy in four main steps, as follows.

First, ICT literacy was seen as measurement of a discrete set of knowledge about computers and their use, coalescing into the concept of ICT literacy in the early years of the field. Second, ICT literacy was seen as measuring a broad set of skills that have links to many traditional and non-traditional school subjects, incorporating the move to technology integration in education. Third, ICT literacy was seen as measuring a set of *progress variables*, which are essential tools for the design of curriculum and assessments in terms of competencies. The "competencies" view depicts the need to understand initial ICT knowledge likely to emerge followed by a developing picture of mastery. In this perspective, students are acknowledged as not being one-size-fits-all in ICT literacy but moving toward increasing competency in their virtual skills, knowledge, competency, awareness, and use. Fourth, the measurement is seen as needing to reflect the "network" perspective on ICT–the critical need for building the power of virtual skills through proficiency with networks of people, information, tools, and resources. A new framework for assessing student ICT learning, based on a learning progression and network point of view has been offered in Wilson & Scalise (2014). This paper goes beyond that introduction and presents an initial analysis of the results from a pilot test looking across the multiple dimensions of the construct.

Methods

The ATC21S methodology group has described that, in order to achieve a working hypothesis of such a complex domain, one *a priori* approach by which known information is incorporated is to describe "dimensions of progression," or theoretical maps of intended constructs, in terms of depth, breadth and how the skills change as they mature for students (Wilson, et al., 2012). For this, the ATC21S project set up an expert panel of ICT experts who turned to the research literature to inform expanded definitions of digital literacy.

The ATC21S panel of ICT experts looked into a wide range of topics in laying out the background for their work in developing the learning progression, including in the areas of augmented social cognition (Chi et al., 2008), applied cognition (Rogers, Pak, & Fisk, 2007), team cognition (Cooke, Gorman, & Winner, 2007), social participation (Dhar & Olson, 1989), cognitive models of human-information interaction (Pirolli, 2007), technological support for work group collaboration (Lampe et al., 2010; Pirolli, Preece, & Shneiderman, 2010), theories of measurement for modeling individual and collective cognition in social systems (Pirolli & Wilson, 1998), topics in semantic representation (Griffiths & Steyvers, 2007), and social information foraging models (Pirolli, 2009).

To take just one example, the research in augmented social cognition field (Chi et al., 2008) describes the emergence of the ability of a group of people to remember, think and reason together. This field investigates how people augment their speed and capacity to acquire, produce, communicate and use knowledge, and to advance collective and individual intelligence in socially mediated environments. This is seen as very relevant to the ICT field, as it is expected that augmented digital and virtual settings will be increasingly common learning environments for students in the 21st Century.

The ATC21S panel of experts identified a set of distinctive ICT literacy goals for students. Earlier frameworks had included individual consumer skills, often on a Web 1.0 model of data repositories that could be accessed over the internet by students. A variety of producer skills, in which students needed to manipulate digital assets in new ways—due to the emergence of Web 2.0 technologies—were seen as emerging trends. Lastly, they found that, as described in documents from the National Initiative for Social Participation (NISP) (e.g., Lampe et al., 2010; Pirolli, Preece, and Shneiderman, 2010), the field of ICT literacy was on the cusp of recognizing the importance of digital networks to education, which would require from the students both "social capital" skills and the ability to draw on the "intellectual capital" of groups and teams. Included in intellectual capital are the Web 3.0 skills of "semantics," or meaning-making through technology, including complex tools such as analytics, the effective use and evaluation of ratings, crowd-sourcing, peer evaluation, tagging, and the ability to judge the credibility and viability of sources.

The expert panel then challenged itself to define, for each of the competencies within the learning progression, increasing levels of sophistication of student competence, that is, to describe what having "more" and "less" of the competency would look like, for age 11-, 13- and 15-year-old students. In other words, as one expert noted, "When someone gets better at it, what are they getting better at?" This might also be thought of as describing how the field of education will document the range of competencies that we might expect as students develop in sophistication.

The BEAR Assessment System

The BEAR Assessment System (BAS) consists of employing data analytics structured around four basic principles of assessment, and four accompanying building blocks that are tools to help develop the assessment. The principles and building blocks are as follows:

> Principle 1: Assessment should be based on a developmental perspective of student learning; the building block is a *construct map* of a progress variable that visualizes how students develop and how we think about their possible changes in response to items. Data analytics in this case are structured around the theoretical conception of the construct map.
>
> Principle 2: There must be a match between what is taught and what is assessed; the building block is the *items design*, which describes the most important features of the format of the items—the central issue, though, is how the items design results in responses that can be analytically related back to the levels of the construct map.
>
> Principle 3: Teachers must be the managers of the system, with the tools to use it efficiently and

effectively; the building block is the *outcome space*, or the set of categories of student responses that make sense to teachers. These categories of student responses become the core of quantitative measures for conclusions through the data analytics.

Principle 4: There is evidence of quality in terms of reliability and validity studies and evidence of fairness, through the data analytics; the building block is an algorithm specified as a *measurement model* that provides for a visual representation of the students and the items on the same graph (called a "Wright Map"), and a number of other data analytic tools that are helpful for testing the quality of the measurement. (Wilson, 2005)

How these principles become embedded in the process and the product of the assessment development, and the nature of these building blocks, is exemplified in the account below.

The Construct Maps. Levels in the ATC21S Learning in Digital Communities learning progression framework, as described below, were developed using the BAS approach described in the previous section, and draw substantively from the literature noted above. The levels in each strand follow a similar valuing, starting first with awareness and basic use of tools, and then building to more complex applications. Evaluative and judgmental skills improve. Leadership and the ability to manage and create new approaches emerge as the more foundational skills are mastered.

The skills can be seen as building mastery, in a theoretical sense, but more information from student work is needed to help validate frameworks. Examples of student work can be collected in cognitive laboratories as was done here or in other venues. Such frameworks become especially useful if ranges of performances can be collected that successfully populate the strands and levels of expected performances. Thus a working definition of the competencies is strengthened through exemplification.

Further large-scale data collection and analysis of results also helps to validate, or refine and reconfigure meaningful frameworks. An important consideration in emerging areas such as these is not only what student performances are currently, but what teachers, schools and states would like them to be. Validated working definitions help school systems plan instructional interventions and assessments, and work with teachers for professional development.

The Four Strands. For the ATC21S project effort, ultimately the focus of ICT Literacy was on *learning in networks*, which was seen as being made up of four strands of a learning progression:

- Functioning as a consumer in network;
- Functioning as a producer in networks;
- Participating in the development of social capital through networks;
- Participating in intellectual capital (i.e., collective intelligence) in networks.

The four strands are seen as interacting together in the activity of learning in networks. They are conceptualized as parallel developments that are interconnected and make up that part of ICT literacy that is concerned with learning in networks.

First, functioning as a Consumer in Networks (CiN) involves obtaining, managing and utilizing information and knowledge from shared digital resources and experts in order to benefit private and professional lives. It involves questions such as:

- Will a user be able to ascertain how to perform tasks (e.g. by exploration of the interface) without explicit instruction?

- How efficiently does an experienced user use a device, application, or other ICT strategy to find answers to a question?
- What arrangement of information on a display yields more effective visual search?
- How difficult will it be for a user to find information on a website?

Second, functioning as a Producer in Networks (PiN) involves creating, developing, organizing and re-organizing information/knowledge in order to contribute to shared digital resources.

Third, developing and sustaining Social Capital through Networks (SCN) involves using, developing, moderating, leading and brokering the connections within and between individuals and social groups in order to marshal collaborative action, build communities, maintain an awareness of opportunities and integrate diverse perspectives at community, societal and global levels.

Fourth, developing and sustaining Intellectual Capital through Networks (ICN) involves understanding how tools, media, and social networks operate, and using appropriate techniques through these resources to build collective intelligence and integrate new insights into personal understandings. An initial idea of the relationship amongst these four strands of the learning progression is given in Figure 2. The four columns are the four strands, each with defined levels within them (exemplified in the next paragraph), and the arrows show expected tracks of development among these constructs. An example of an activity that is hypothesized to map into the ICN2 level ("Functional builder") of the Intellectual Capital strand is demonstrated in the chat log in Figure 3. In this example, students are creatively using tags to make their communication clearer. Specifically, the team members have devised a strategy using different fonts and/or colors to distinguish what each team member is typing.

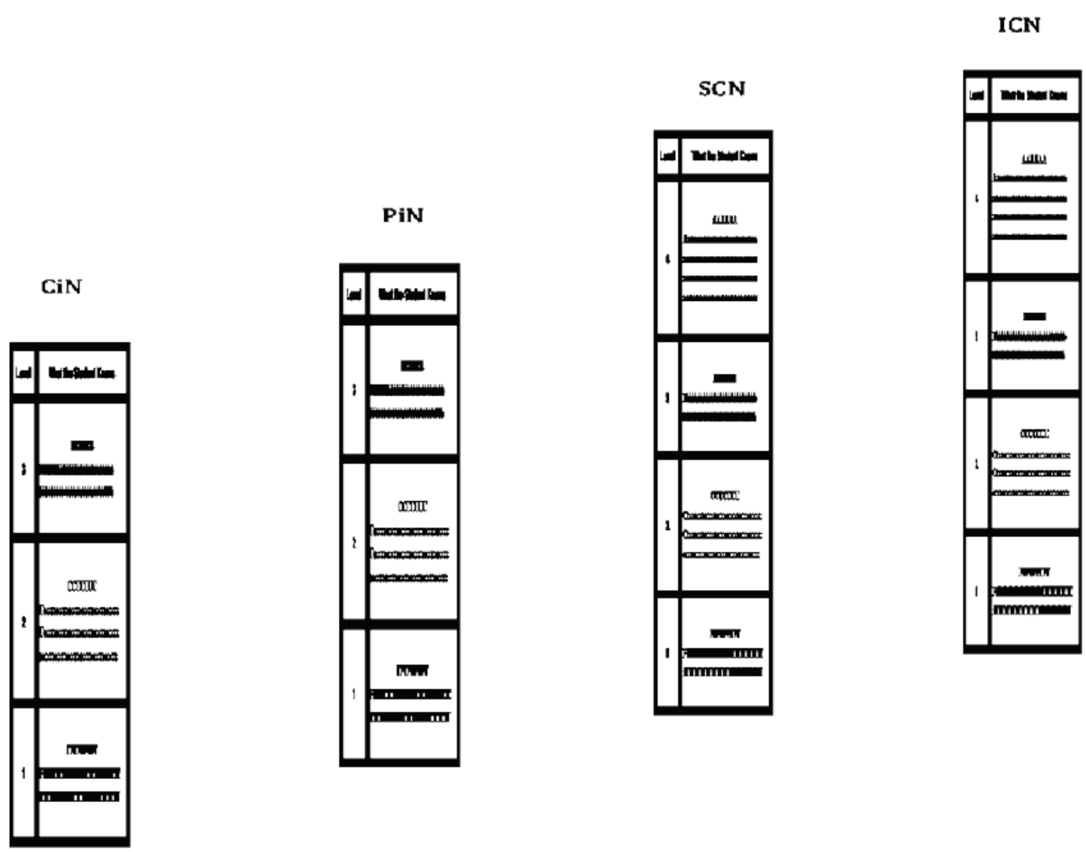


Figure 2. The four strands of ICT Literacy, represented as a four-part learning progression.

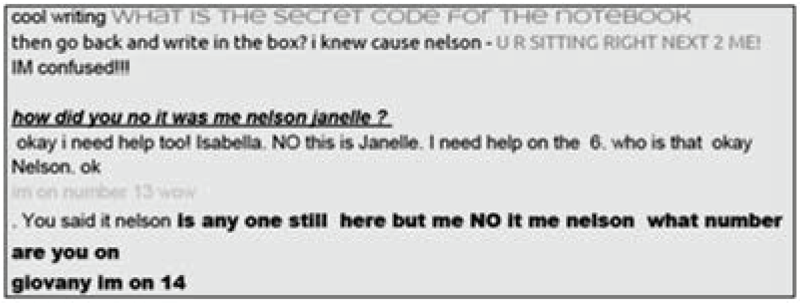


Figure 3. Students' chat log.

In Table 1 the levels of the second of these four strands have been described as a hypothesized construct map showing an ordering of skills or competencies involved in each. At the lowest levels are the competencies that one would expect to see exhibited by a novice or beginner. At the top of the table are the competencies that one would expect to see exhibited by an experienced person – someone who would be considered very highly literate in ICT. The construct map is hierarchical in the sense that a person who would normally exhibit competencies at a higher level would also be expected to be able exhibit the competencies at lower levels of the hierarchy. The maps are also probabilistic in the sense that they represent different probabilities that a given competence would be expected to be exhibited in a particular context rather than certainties that the competence would always be exhibited. For a complete account of the construct maps for the four competencies, see Wilson & Scalise (2014).

Table 1. Functioning as a Producer in Networks (PiN)

	PRODUCER IN NETWORKS
	Creative producer
PiN3	Team situational awareness in process Optimize assembly of distributed contribution to products Extending advanced models (e.g. business models) Producing attractive digital products using multiple technologies / tools Choosing among technological options for producing digital products
	Functional producer
PiN2	Establishing and managing networks & communities Awareness of planning for building attractive websites, blogs, games Organizing communication within social networks Developing models based on established knowledge Developing creative & expressive content artifacts Awareness of security & safety issues (ethical and legal aspects) Using networking tools and styles for communication among people
	Emerging producer
PiN1	Produce simple representations from templates Start an identity Use a computer interface Post an artifact

The Items Design. The Berkeley Evaluation and Assessment Research (BEAR) Center at UC Berkeley developed three scenarios in which to place tasks and questions that could be used as items to indicate where a student might be placed along each of the four strands.

Observation formats included numerous innovative types (Scalise & Gifford, 2006). The data set included process data from activities in the tasks, collaboration data from chat logs and other activities, explanations and argumentation from students, and answers to technology-enhanced selected response queries of various types. Some examples are illustrated in the upcoming figures, when each task is introduced.

Each scenario was designed to address more than one strand of the construct, but there were different emphases in how the strands area were represented among the scenarios. Where possible, we took advantage of existing web-based tools for instructional development. These are each briefly described below.

Arctic Trek. One potential mechanism for the assessment of student ability in the learning network aspect of ICT literacy is to model assessment practice through a set of exemplary classroom materials. The module that has been developed is based on the Go North/Polar Husky information website (www.polarhusky.com) run by the University of Minnesota. The Go North website is an online adventure learning project based around arctic environmental expeditions. The website is a learning hub with a broad range of information and many different mechanisms to support networking with students, teachers and experts. ICT literacy resources developed relating to this module focus mainly on the functioning as a Consumer in Networks strand. The tour through the site for the ATC21S demonstration scenario is conceived as a "collaboration contest," or virtual treasure hunt (see Figure 4 for a sample screen). The Arctic Trek scenario views social networks through ICT as an aggregation of different tools, resources and people that together build community in areas of interest. In this task, students in small teams ponder tools and approaches to unravel clues through the Go North site, via touring scientific and mathematics expeditions of actual scientists. The Arctic Trek task, in which students work in teams, is demonstrated in Figure 5. In that task, students are expected to find the colors that are used to describe the bear population in the table, part of which is shown at the top. The highlighted chat log of students at the bottom of the figure indicates that students are communicating in order to identify signal versus noise in the supplied information. The colors in the text are the colors shown in the columns on the right of the table. Requiring both identifying signal versus noise in information and interrogating data for meaning, this performance can be mapped into the ICN3 level ("Proficient builder") of the ICN strand (Wilson & Scalise, 2014). For further examples of activities and items from the Arctic Trek scenario, see Wilson & Scalise (2014).

Figure 4. A sample screen from the Arctic Trek task: the Welcome screen.

Norwegian Bay	190 (1998)	102-278					4	4	Data deficient	Declining	Very high
Southern Beaufort Sea	1526 (2006)	1210-1842					44	80	Reduced	Declining	Moderate
Southern Hudson Bay	900-1000 (2005)	396-950 (ON) 70-100 (James Bay)					35	61	Not reduced	Stable	Very high

Clue 3
Find the table that tells how many there are of me. How many colors are used to describe the bear population? (use status, trend and risk columns)
Answer:I looked and counted the colors 5 -Alex, Team Captain
I think there should be six because the “data deficient” is also a color. -Andy

Figure 5. Example of students’ work analyzing the table.

Webspiration. In the second demonstration task, framed as part of a poetry work unit, students of ages 11-15 read and analyze well-known poems. In a typical school context, we might imagine that the teacher has noticed that his or her students are having difficulty articulating the moods and meanings of some of the poems. In traditional teacher-centered instruction regarding literature, the student role tends to be passive. Often, teachers find that students are not spontaneous in their responses to the poems, but may tend to wait to hear what the teacher has to say, and then agree with what is said. To help encourage students to formulate their own ideas on the poems, we use a collaborative graphic organizer through the Webspiration online tool. The teacher directs the students to use Webspiration to create an idea map collaboratively using the graphic organizer tools, and to analyze each poem they read. Students submit their own ideas and/or build on classmates’ thoughts. For examples of activities and items from the Webspiration scenario, see Wilson & Scalise (2014).

Second Language Chat. This scenario was developed as a peer-based second language learning environment through which students interact in learning. Developing proficiency in a second language (as well as in the mother tongue) requires ample opportunities to read, write, listen and speak. This assessment scenario asks students to set up a technology/network-based chat room, invite participants, and facilitate a chat in two languages. It also involves evaluating the chat and working with virtual rating systems and online tools such as spreadsheets. Worldwide, "conversation partner" language programs such as this have sprung up in recent years. They bring together students wishing to practice a language with native speakers, often in far-flung parts of the world. The cultural and language exchanges that result demonstrate how schools can dissolve the physical boundaries of walls and classrooms. Collaborative activities also tap rich new learning spaces through the communication networks of ICT literacy. This task shows how collaborative activities can also provide ample assessment opportunities in digital literacy. For examples of activities and items from the Second Language Chat scenario, see Wilson & Scalise (2014).

The Outcome Space. Each item that was developed was targeted at one or more of the four strands, and the expected range of levels that would be represented in the item responses were also noted. Where the responses are selected from a fixed set (as in a multiple-choice item), the set can be planned ahead of time, but for open-ended tasks and activities, the resulting work product is more complex. Examining the student work product is a something that needs to be empirically investigated. The tabulation is shown in Figure 6. As can be seen, the first three levels were reasonably well-covered, but level 4, which we seldom expect to see for students in this population, had only one instance.

Levels (Progressive)	ICT Literacy—Learning in digital communities CONSTRUCT/Learning Outcomes				Total
	Consumer	Producer	Social Capital	Intellectual Capital	
Level 4	N/A	N/A	Web 0 Arctic 1 2LChat 0	Web 0 Arctic 0 2LChat 0	Web: 0 Arctic: 1 2LChat: 0
Level 3	Web 0 Arctic 2 2LChat 0	Web 0 Arctic 2 2LChat 0	Web 0 Arctic 6 2LChat 1	Web 10 Arctic 2 2LChat 1	Web: 10 Arctic: 12 2LChat: 2
Level 2	Web 8 Arctic 6 2LChat 0	Web 4 Arctic 16 2LChat 8	Web 7 Arctic 0 2LChat 6	Web 6 Arctic 7 2LChat 0	Web: 25 Arctic: 29 2LChat: 14
Level 1	Web 2 Arctic 2 2LChat 2	Web 4 Arctic 0 2LChat 6	Web 1 Arctic 0 2LChat 6	Web 2 Arctic 2 2LChat 0	Web: 9 Arctic: 4 2LChat: 14
Total	Web: 10 Arctic: 10 2LChat: 2	Web: 8 Arctic: 18 2LChat: 14	Web: 8 Arctic: 7 2LChat: 13	Web: 18 Arctic: 11 2LChat: 1	120 per age group plus PU/FP/HUE, Practice, & Covariates

***Some CR items (Constructed Response) will measure *up through* the listed level (listed level is top score).**

Figure 6. The number of data points from each scenario, and their planned allocation to the levels from each strand.

The Measurement Model: Algorithm Used for Learning Analytics. In this approach to data analytics, a statistical analytic technique called a "measurement model" serves as the algorithm to compile the results and make inferences about learners. Other fields such as computer science that come to learning analytics from a different historical basis often use different vocabulary to describe such algorithms. For instance, the Rasch model often used in educational assessment from a computer science perspective would be considered an algorithm employing a multilayer feed-forward network (Russell & Norvig, 2009) with *g* as the Rasch function (a semi-linear or sigmoidal curve-fitting function), in which weights (item discrimination) are constrained to one for all inputs, and the item parameters estimated are only the thresholds on each item node (item difficulty). The 2PL IRT model, by contrast, is an algorithm employing a multilayer feed-forward network with *g* as the 2PL function (also a sigmoidal curve-fitting function), in which both weights (item discrimination) and thresholds on each item node (item difficulty) are estimated. The 3PL model is an algorithm employing a multilayer feed-forward network with *g* as the 3PL function (sigmoidal), in which weights (item discrimination), thresholds on each item node (item difficulty), and a lower asymptote (guessing parameter) are estimated.

For this initial data collection, we chose to use the Random Coefficients Multinomial Logit Model (Adams & Wilson, 1996), which allows polytomous data (as is appropriate for the items in the tasks), and provides the Wright Maps mentioned above (and are shown in Figures 7, 8 and 10 below). Due to small sample size, we ended up with only two scored items in the SCN strand and therefore excluded this strand from the analyses. As a result, we report our findings from three separate unidimensional analyses with 27 items measuring the CiN strand, nine items measuring the PiN strand,

and 22 items measuring the ICN strand. The ConQuest 3.0 estimation software (Adams, Wu & Wilson, 2012) was employed throughout the analyses, which uses RCML as its basis for statistical modeling. See the paper by Adams et al. (1997) for the relevant equations and statistical estimation procedures for the analysis.

Results and Discussion: Coming to an Interpretation

Two of the three scenarios were selected for empirical study with students. The two, the science/math Arctic Trek collaboration contest and the Webspiration shared literature analysis task, were identified by participating countries as the most desirable to study at this time. The science/math and language arts activities were more aligned with the school systems in the countries, which rarely used anything like cross-country chat tools in the classroom, but sometimes did employ math simulations and online scientific documents as well as graphical and drawing tools for student use. In particular, the third task (the Second Language Chat) was described by participating countries, teachers, and schools as a forward-looking, intriguing scenario, but farther away on the adoption curve for school-based technology.

Not all of the planned automated scoring and data analysis for the items in the two piloted scenarios was available for application to this data set, as the total numbers of cases was too small for the empirically-based scoring to be successfully calibrated. This calibration? will be completed at a later point, when larger data sets are available. Each of the two scenarios were presented in three forms, for 11-, 13- and 15 year-olds, respectively, with a subset of common items across the three forms.

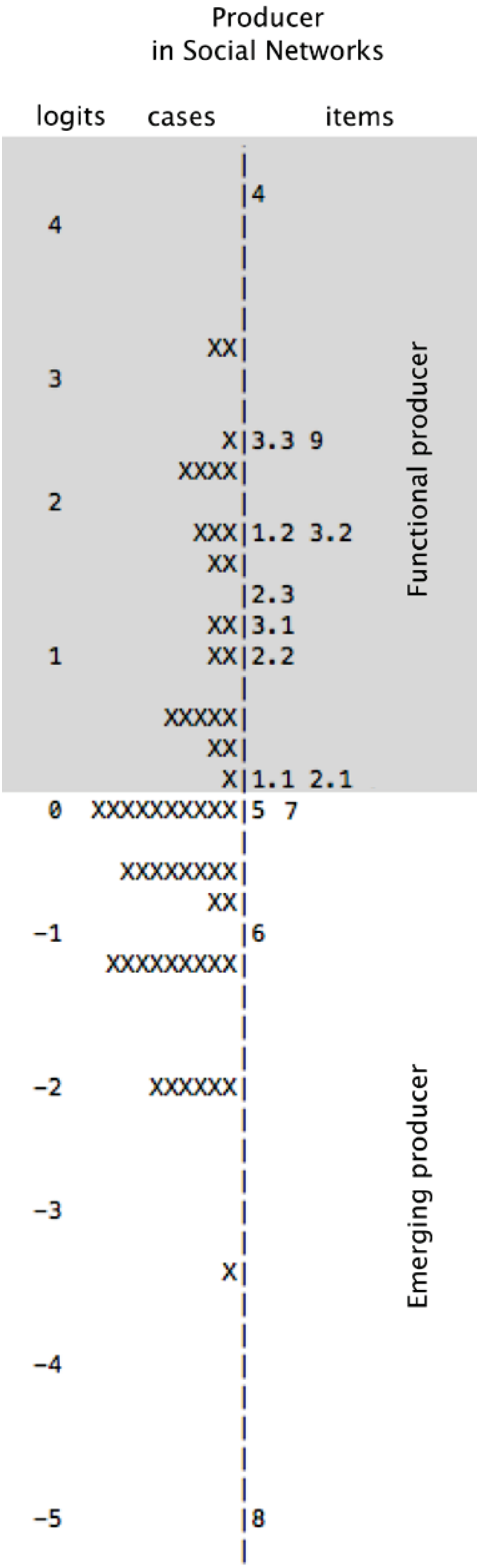
For a sample of 103 students in our first field test (i.e., those for whom we were able to match their login ids for two scenarios), assessment results from the two scenarios within the overall ICT literacy domain were analyzed using a three separate *consecutive* unidimensional item response (RCML) models.

One way of checking if the assumptions and requirements of the confirmatory approach has been met within each of the three consecutive models is to examine the weighted mean square fit statistic estimated for each item within each model. Item fit can be seen as a measure of the discrepancy between the observed item characteristic curve and the theoretical item characteristic curve (Wu & Adams, 2010). ConQuest estimates the residual-based weighted fit statistics, also called infit; ideally, infit values are expected to be close to 1.0. Values less than 1.0 imply that the observed variance is less than the expected variance while values more than 1.0 imply that the observed variance is more than the expected variance. A common convention of 3/4 (0.75) and 4/3 (1.33) is used as acceptable lower and upper bounds (Adams & Khoo, 1996). After deleting items that fit poorly, we found that all of the remaining 58 items for the three consecutive models fell within this range.

Table 2 shows the variances and correlations of EAP scores obtained from the consecutive approach. The closest estimated relation is between the CiN and PiN dimensions, at a correlation of .69.

Table 2. Variances and correlations of EAP scores from three consecutive unidimensional models.

	CiN	PiN	ICN
PiN	0.65		
ICN	0.57	0.57	
variance	0.97 (0.14)	2.54 (0.35)	1.79 (0.25)



Reliabilities for each of the three dimensions are .76, .79, .73, for CiN, PiN, and ICN, respectively—all of which are quite high, given the number of items available.

<Figure 7. Wright Map for the Consumer in Social Networks strand with bands representing each level.

Figure 7 shows one of the “Wright Maps” (Wilson, 2005) obtained from the unidimensional model for Consumer in Social Networks. Items are vertically ordered with respect to their difficulties and persons (cases) are vertically ordered with respect to their abilities. Each “X” on the left-hand side represents a small number of students, and the items are shown on the right-hand side using their item numbers. The locations are interpreted as follows, for dichotomous items:

(a) when a student’s X matches an item location, the probability of that student succeeding on that item is expected to be 0.50;
(b) when a student’s X is above the item, then the probability is above 0.5 (and vice-versa); and
(c) these probabilities are governed by a logistic distribution (see Wilson, 2005 for a discussion of this).

Where the items are polytomous, the labeling is a little more complex, for example, in Figure 7, note that Item 5 is represented by two labels: 5.1 and 5.2. The former is used to indicate the threshold between category 0 and categories 1 and 2 (combined); the latter is used to represent the threshold between categories 0 and 1 (combined) and category 2. The interpretation of the probability is equivalent to that for a dichotomous item: that is, when a student’s X matches the 5.1 location, the probability of that student succeeding at levels 1 or 2 on that item is expected to be 0.50; and similarly, when a student’s X matches the 5.2 location, the probability of that student succeeding at only level 2 on that item is expected to be 0.50.

Just as schools have offered a more level playing field for the learning of traditional knowledge areas such as math and reading, so too direct intervention may be needed to bring all students to career and college readiness skills in their digital undertakings. Other 21st century skills should be similarly assessed. Given the importance to the future opportunities that students will have depending on the types of skills and abilities they can exhibit in these areas, the contributions they will need to make through the use of such skills cannot be underestimated (Binkley et al., 2012). Again, more examples that show the possibility of more fine-grain interpretation at the individual level through such learning analytic approaches will be discussed in an upcoming paper.

The maps are also useful in investigating whether there is a good coverage of abilities by items. Ideally, if permitted a sufficient number of items for each strand, we would hope to see the range of item difficulties approximately match the range of person abilities. This would mean that there are items approximately matching every level of the person ability. This can be seen to be true for the Consumer in Social Networks strand. Figure 7 also shows "banding" for the Wright Map for the Consumer in Social Networks—that is, we have carried out a judgmental exercise to approximately locate where students in each level would be located. We note that, after misfit analysis, only level 3 of item 7 (i.e., 7.3) remains in the item pool to represent the highest level, the discriminating consumer. In contrast, the lower two levels, conscious consumer and emerging consumer, are well-represented. We can see that students in this sample have shown a range of abilities on this strand that spans all three of the hypothesized levels, from Emerging to Discriminating Consumer, although, as we might have expected, there are relatively more in the lower levels than in the higher levels (an observation that will be repeated in the other strands).

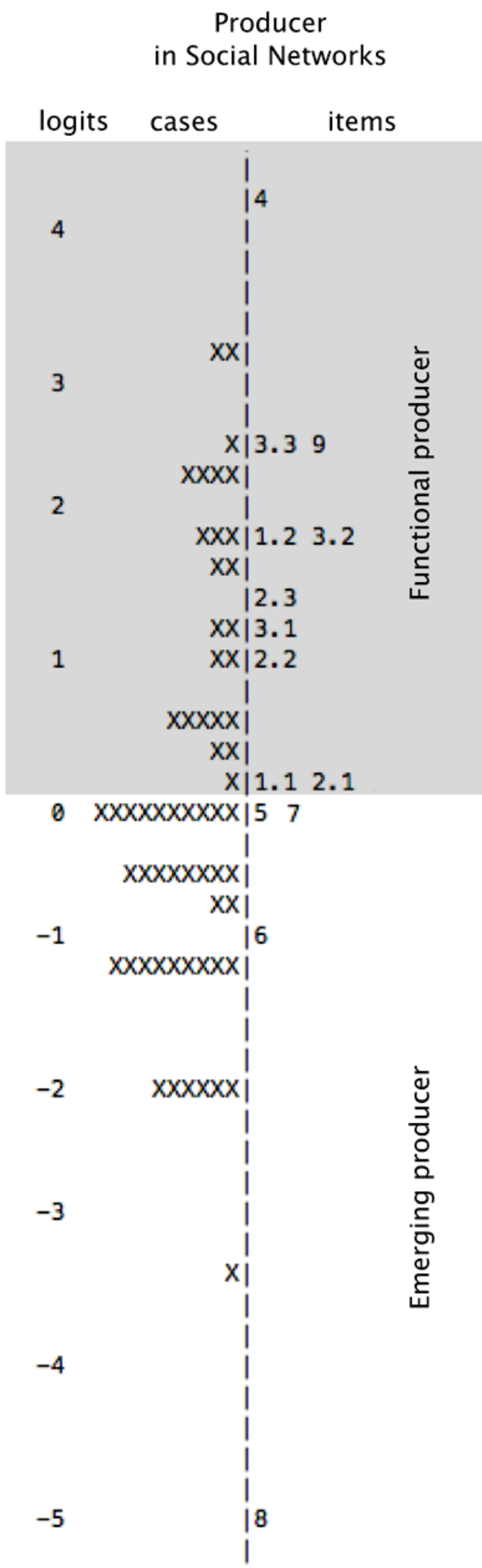
Interpretive Phase

Once the coherence of the four building blocks is in place and results are obtained, interpretation of the learning analytics can take place. This interpretation is done through the maps yielded. For example, the distribution of students over the proposed construct for the Consumer strand can be seen in Figure 7. Most students fell into the Conscious consumer portion of the construct, as shown by the histogram composed of the "X" character on the left portion of the display. This location on the map indicated that the students tended to approach the ICT learning in social network activities with the ability to select appropriate tools and strategies (strategic competence), construct targeted searches, compile information systematically, and at least know that credibility of online information is an issue, though perhaps not how to address this. Some students, represented by the lower X's in this band, represented some but not full proficiency in these areas while others, as they placed in locations up the histogram, could be expected to employ these behaviors effectively more often and more fully.

However, a few students remained in the Emerging consumer terrain. Here, students could perform basic searches for information and knew that some digital tools existed to do more. But they were often routinely unable to employ digital tools to perform more than the most basic tasks. Some students were challenged even with the notion of clicking a link to navigate to another page, where the information they sought could be found. Others exhibited only the concept of linear reading, where they would attempt to read through each web page they reached from top to bottom. Without the concept of scanning or searching for information within a page, they were unable to progress very far with the activities, or contribute effectively to collaborations.

Also, students in the Emerging terrain often exhibited no notion of questioning sources or considering the credibility of information. These students would need substantial support with their digital literacy skills to take part in virtual or distributed collaborative learning.

By contrast, a few students showed behaviors across settings that would propel them into the upper category of Discriminating consumer proficiencies. These students were showing leadership ability with their ICT collaboration skills. They could not only question but also judge credibility of sources. They often were able to integrate information into larger views of coherent knowledge, construct searches suited to personal circumstances, and then filter, manage, and organize information much more effectively than other students. Because they had a tendency to seek expert knowledge and select optimal tools, their collaboration skills and their teams benefitted. Groups in which they engaged might take leaps forward based on such participants. A notion of how to combine teams so that they represented a range of support and incorporated peer learning would probably be helpful to support the wide range of skills.



A more fine-grained interpretation at the level of each student can also take place using these maps. This is needed for clear decision-making about current learning outcomes and future needs if intervention will take place at the student level. For this purpose, a student at any location on the map can be selected. Based on their map location, inferences can be made about what they have mastered in this part of the complex terrain of 21st century skills. Such examples at the individual level will be discussed in an upcoming paper.

For the Producer strand of the construct, the group of students surveyed tended to be more evenly split between the Emerging and Functional producer traits. This positioning? indicates that though living in the era of Web 3.0, when students may be immersed in a Maker culture, some students have considerably more skills toward producing digitally and collaboratively online than do other students. About half the students could produce simple representations from templates or start an online identity, but were challenged to go much beyond this to participate in an effective collaboration, problem solve, or produce a product. By contrast, the other approximately half of students were able to go far beyond this. They showed the ability to organize communication effectively in social networks, use collaborative networked tools to develop creative and expressive artifacts, plan for shared solutions, and also demonstrate? an awareness of safety and security online in the process of their production activities.

So the message here, should it be verified by larger data sets, is a strong word of caution to programs, schools or instructors. Educators may presume every K-12 student today is a digital native. It is true that the digital divide may have eased in many locations and that students in a number of primary and secondary classrooms today may have grown up surrounded by many new technologies. However, the matter of access is not the sole determining factor in student development of 21st century skills and abilities (Binkley et al., 2012).

<Figure 8. Wright Map for the Producer in Social Networks strand with bands representing each level.

Figure 8 shows the banded Wright Map for the Producer in Social Networks strand. All of the items in this strand ended up in their hypothesized band. Items in the Functional Producer band require students to develop creative, expressive and/or useful content artifacts and tools. For instance, item 2 (shown in Figure 9 below) and item 3 require students to create a graph representing data. In item 1, we asked students to summarize the data in a table. We hypothesized that all of these three items would end up in the upper band of the Wright Map and our hypothesis was confirmed. The two most difficult items in this band, as we hypothesized, are items 4 and 9. In both of these items we asked students to upload files to the webpage. Items in the Emerging Producer band require students to post an artifact and/or perform basic production tasks. In items 5, 6 and 7, for instance, we asked students to copy/paste the readily available content. In item 8, we asked students to connect at least one node in the concept map. All of these items ended up in the lower part of the Wright Map, as we hypothesized before the test.

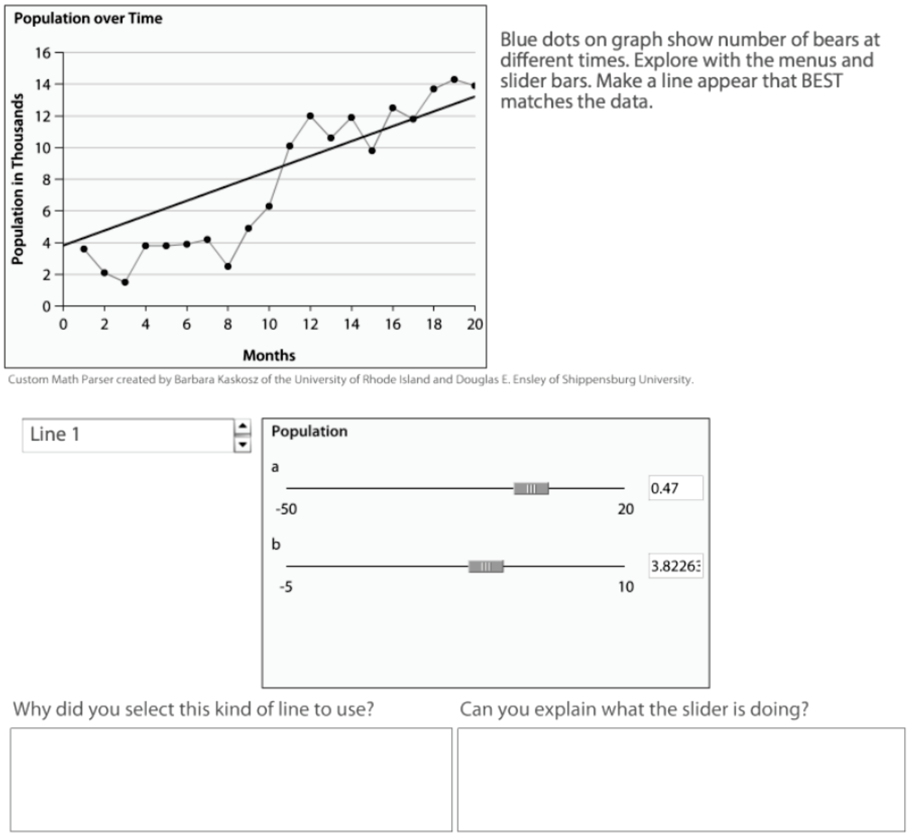


Figure 9. Task for item 2

We do not display a banded Wright Map for the Developer of Social Capital strand, as there are only two items remaining in this strand, both of which are aimed at measuring at third level (Proficient connector) of the construct. Finally, we do not display and discuss the Wright Map results for the Participator in the Intellectual Capital strand, as it will appear in a future paper.

Conclusion

In conclusion, this paper suggests a modified definition for learning analytics. The definition underscores the need for meaningful interpretation of learning results and also incorporates key ways that learning analytics are employed with complex data and algorithms.

Measuring intricate constructs through digital analytics, as displayed here, can help us to appreciate the new ways that students will be required to think and work compared to previous decades. This paper describes a domain modeling process for assessing ICT literacy through the BEAR

Assessment System, along with examples of task development, and results from implementation of a first field test in four countries.

This initial field test has demonstrated through the learning analytics approach that, for this data set, students in the 11-15-year age group show widely differing knowledge and skills in these areas. Some students are only just beginning to take initial steps toward digital competence while other students show levels of mastery that would likely challenge skilled adult samples, such as collaborating easily to create professional-level commentaries or reports. Doubtless the wide range is due at least in part to the absence of formal teaching and opportunities to learn these skills in many schools, which results in an increasing gap between the haves and the have-nots. It looks as though this is particularly true in the domain of ICT literacy, which we take to be a gatekeeper for other skills and achievements in the future.

While these techniques and approaches are illustrated through one example project, they potentially are broadly applicable to many learning domains. The takeaway lesson may be that effectively educating students for 21st century skills may also require offering them 21st century supports for learning—including improved approaches to examining patterns and evidence of learning in complex settings.

References

ACARA. (2012). Australian Curriculum, Assessment and Reporting Authority, National Assessment Program, ICT Literacy, Years 6 and 10 Report 2011.

Adams R. J., & Khoo, S-T. (1996). *Quest* [computer program]. Melbourne, Australia: ACER Press.

Adams, R.J., & Wilson, M. (1996). Formulating the Rasch model as a mixed coefficients multinomial logit. In, G. Engelhard & M. Wilson (Eds.), *Objective measurement III: Theory into practice*. Norwood, NJ: Ablex.

Adams, R.J., Wu, M., & Wilson, M. (2012). *ConQuest 3.0* [computer program]. Hawthorn, Australia: ACER.

Ainley, J., Fraillon, J., Schulz, W., & Gebhardt, E. (2014). *Measuring Changes in ICT Literacy over Time*. Paper presented at the National Council on Measurement in Education, Philadelphia, PA.

Baker, R. S., & Siemens, G. (2014). Educational data mining and learning analytics. In K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 253-274).

Binkley, M., Erstad, O, Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining Twenty-First century skills. In Griffin, P, McGaw B. & Care, E. (Eds.), *Assessment and Teaching of 21st Century Skills*. Dordrecht, Germany: Springer.

Brady, A., Conlan, O., Wade, V., & Dagger, D. (2006). *Supporting Users in Creating Pedagogically Sound Personalised Learning Objects.* Paper presented at the Adaptive Hypermedia and Adaptive Web-based Systems, Dublin, Ireland.

Briggs, D. C., & Wilson, M. (2003). An introduction to multidimensional measurement using Rasch models. *Journal of Applied Measurement*, *4*(1), 87–100.

Chedrawy, Z., & Abidi, S. S. R. (2006). *An Adaptive Personalized Recommendation Strategy Featuring Context Sensitive Content Adaptation.* Paper presented at the Adaptive Hypermedia and Adaptive Web-Based Systems, 4th International Conference, AH 2006, Dublin, Ireland.

Chi, E. H., Pirolli, P., Suh, B., Kittur, A., Pendleton, B., & Mytkowicz, T. (2008). Augmented Social Cognition. Palo Alto, CA: Palo Alto Research Center.

Cisco, Intel, Microsoft (CIM). (2008). *Transforming Education: Assessing and Teaching 21st Century Skills.* Authors. Retrieved from http://atc21s.org/wp-content/uploads/2011/04/Cisco-Intel-Microsoft-Assessment-Call-to-Action.pdf

Cooke, N. J., Gorman, J. C., & Winner, J. L. (2007). Team Cognition. In F. T. Durso (Ed.), *Handbook of Applied Cognition* (pp. 239-268). New York: John Wiley & Sons, Ltd.

Dagger, D., Wade, V., & Conlan, O. (2005). Personalisation for All: Making Adaptive Course Composition Easy. *Educational Technology & Society, 8*(3), 9-25.

Dhar, V., & Olson, M. H. (1989). Assumptions Underlying Systems that Support Work Group Collaboration. In M. H. Olson (Ed.), *Technology Support for Work Group Collaboration*. Hillsdate, NJ: Lawrence Erlbaum Associates.

Dringus, L. P. (2012). Learning analytics considered harmful. *Journal of Asynchronous Learning Networks, 16*(3), 87-100.

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning, 4*(5/6), 301-317. doi: 10.1504/IJTEL.2012.051816

Ferrari, A. (2012). *Digital competence in practice: An analysis of frameworks*. Seville, Spain: Institute for Prospective Technological Studies, European Commission.

Fraillon, J. (2014). *Assessing Computer and Information Literacy across 20 Countries: The International Computer and Information Literacy (ICILS)*. Paper presented at the National Council on Measurement in Education (NCME), Philadelphia, PA.

Friedman, T. L. (2007). *The World is Flat 3.0: A Brief History of the Twenty-first Century*. New York: Picador.

Griffin, P., McGaw B. & Care, E. (Eds.) (2012). *Assessment and Teaching of 21st Century Skills*. Dordrecht, Germany: Springer.

Griffiths, T. L., & Steyvers, M. (2007). Topics in Semantic Representation. *Psychological Review, 114*(2), 211-244.

Kennedy, C. A., & Draney, K. (2006). Interpreting and using multidimensional performance data to improve learning. In X. Liu (Ed.), *Applications of Rasch Measurement to Science Education*. Chicago: JAM Press.

Kozma, R. (Ed.). (2003). *Technology, innovation, and educational change. A global perspective. A report of the Second Information Technology in Education Study, Module 2*. Amsterdam: International Association for the Evaluation of Educational Achievement.

Kozma, R. B. (2003). Technology and classroom practices: An international study. *Journal of Research on Technology in Education, 36*(1), 1-5.

Lampe, C., Resnick, P., Forte, A., Yardi, S., Rotman, D., Marshall, T. & Lutters, W. (2010). Educational priorities for technology-mediated social participation. *IEEE Computer*, *43*(*11*), 60-67.

Lehrer, R., Kim, M-J., Ayers, E., & Wilson, M. (2014). Toward establishing a learning progression to support the development of statistical reasoning. In J. Confrey and A. Maloney (Eds.), *Learning over Time: Learning Trajectories in Mathematics Education*. Charlotte, NC: Information Age Publishers.

Masters, G. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*(2), 149-174.

National Research Council (NRC). (2001). *Knowing what students know: The science and design of educational assessment.* Committee on the Foundations of Assessment. J. Pellegrino, N. Chudowsky, & R. Glaser (Eds). Washington, DC: National Academy Press.

Papamitsiou, Z., & Economides, A. A. (2014). Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence. *Educational Technology & Society, 17*(4), 49-64.

Park, J. (2011). ICT in Education in Asia-Pacific. *Asia-Pacific Regional Forum on ICT Applications*. Retrieved from http://www.itu.int/ITU-D/asp/CMS/Events/2011/ict-apps/s5_JPark.pdf

Pedersen, S. G., Malmberg, P., Christensen, A. J., Pedersen, M., Nipper, S., Græm, C. D., & Norrgård, J. (Eds.). (2006). *E-learning Nordic 2006: The impact of ICT on educatio*. Copenhagen, Denmark: Ramboll Management.

Pirolli, P. (2007). Cognitive Models of Human-Information Interaction. In F. T. Durso (Ed.), *Handbook of Applied Cognition* (pp. 443-470). New York: John Wiley & Sons, Ltd.

Pirolli, P. (2009). An elementary social information foraging model. *CHI 2009*, *ACM Conference on Human Factors in Computing Systems,* (pp. 605-614). Boston: Association for Computing Machinery.

Pirolli, P., Preece, J., & Shneiderman, B. (2010). Cyberinfrastructure for social action on national priorities. *IEEE Computer, 43*(11), 20-21.

Pirolli, P., & Wilson, M. (1998). A Theory of the Measurement of Knowledge Content, Access, and Learning. *Psychological Review, 105*(1), 58-82.

Qualifications and Curriculum Authority. (2007). *About information and communication technology: Assessment guidance* Retrieved from http://web.archive.org/web/20040815054520/http://www.qca.org.uk/7889.html

Rogers, W. A., Pak, R., & Fisk, A. D. (2007). Applied Cognitive Psychology in the Context of Everyday Living. In F. T. Durso (Ed.), *Handbook of Applied Cognition* (pp. 3-27). New York: John Wiley & Sons, Ltd.

Russell, S., & Norvig, P. (2009). *Artificial Intelligence, A Modern Approach* (3rd Ed.). Upper Saddle River: Prentice Hall.

Scalise, K. (2012). Using Technology to Assess Hard-to-Measure Constructs in the CCSS and to Expand Accessibility. *Invitational Research Symposium on Technology Enhanced Assessments, Educational Testing Service (ETS), Washington, DC*. Retrieved from http://www.k12center.org/events/research_meetings/tea.html

Scalise, K., & Gifford, B. R. (2006). Computer-Based Assessment in E-Learning: A Framework for Constructing "Intermediate Constraint" Questions and Tasks for Technology Platforms. *Journal of Teaching, Learning and Assessment, 4*(6).

Sclater, N. (2014). JISC: Code of practice for learning analytics: A literature review of the ethical and legal issues. Retrieved from http://repository.jisc.ac.uk/5661/1/Learning_Analytics_A-_Literature_Review.pdf

Stanton, J. M. (2012). *An Introduction to Data Science.* Retrieved from http://surface.syr.edu/istpub/165/

U.S. Department of Education. (2010). *Transforming American Education: Learning Powered by Technology. National Education Technology Plan 2010.* Retrieved from http://www.ed.gov/sites/default/files/netp2010.pdf

Wilson, M. (2005). *Constructing Measures: An Item Response Modeling Approach*. Mahwah, NJ: Erlbaum.

Wilson, (2009a, December). *Assessment for Learning* and *for Accountability*. Paper presented at the Exploratory Seminar: Measurement Challenges Within the Race to the Top Agenda, at ETS, Princeton, NJ. Retrieved from

Wilson, (2009b, December). *Assessment for Learning* and *for Accountability*. Policy Brief from the Exploratory Seminar: Measurement Challenges Within the Race to the Top Agenda, at ETS, Princeton, NJ. Retrieved from

Wilson, M. (2012). Responding to a challenge that learning progressions pose to measurement practice: hypothesized links between dimensions of the outcome progression. In A. C. Alonzo & A. W. Gotwals, (Eds.), *Learning Progressions in Science*. Rotterdam, The Netherlands: Sense Publishers.

Wilson, M., Bejar, I., Scalise, K., Templin, J., Wiliam, D., & Torres Irribarra, D. (2010). Assessment and Teaching of 21st Century Skills: Perspectives on Methodological Issues *Learning and Technology World Forum 2010*. White Paper presented at the Learning and Technology World Forum 2010, London.

Wilson, M., Bejar, I., Scalise, K., Templin, J., Wiliam, D., & Torres-Irribarra, D. (2012). Perspectives on methodological issues. In Griffin, P., McGaw B. & Care, E. (Eds.), *Assessment and Teaching of 21st Century Skills*. Dordrecht, Germany: Springer.

Wilson, M. & Carstensen, C. (2007). Assessment to improve learning in mathematics: The BEAR Assessment system. In A. Schoenfeld (Ed.), *Assessing mathematical proficiency* (pp. 311-332). London: Cambridge University Press.

Wilson, M., & Hoskens, M. (2005). Multidimensional item responses: Multimethod/multi-trait perspectives. In S. Alagumalai, D. D. Curtis, & N. Hungi (Eds.) *Applied Rasch measurement: A*

Book of Exemplars (Papers in honour of John Keeves) (pp. 287-308). Dordrecht, The Netherlands: Springer-Kluwer Academic Publishers.

Wilson, M., & Scalise, K. (2015). Assessment of Learning in Digital Networks. In Griffin, P., & Care, E. (Eds.), *Assessment and Teaching of 21st Century Skills Volume 2 - Methods & Approaches*. Dordrecht: Springer.

Wilson, M., Scalise, K., & Gochyyev, P. (2015). Rethinking ICT Literacy: From Computer Skills to Social Network Settings. *Thinking Skills and Creativity, 18*, 65-80.

Wilson, M. & Sloane, K. (2000). From principles to practice: An embedded assessment system. *Applied Measurement in Education, 13*(2), 181-208. Retrieved from http://www.informaworld.com/smpp/content~content=a783685281~db=all

Using Community College Prior Academic Performance to Predict Re-Enrollment at a Four-Year Online University

Denise Nadasen
University of Maryland University College

Alexandra List
Ball State University

Abstract

Students' re-enrollment in the subsequent semester after their first semester at a four-year institution is a strong predictor of retention and graduation. This is especially true for students who transfer from a community college to a four-year institution because of the many external or non-academic factors influencing a student's decision to re-enroll. This research study examines student learner characteristics and course-taking behaviors at the community college and first-term GPA at a four-year institution to predict the likelihood of re-enrollment for 8,200 students from two community colleges who transferred to an online, public, four-year institution. The logistic regression models showed that gender, age, and first-term GPA at the four-year institution were significant predictors of re-enrollment. These findings contribute to the growing literature on transfer students and may provide researchers and practitioners a greater understanding of how community college factors influence the progression and success for transfer students at four-year institutions.

Introduction

Eighty percent of beginning community college students express an interest in transferring to a four-year university; however, within six years of transfer, only 15% of those students starting at a community college graduate with a four-year credential (Shapiro et al., 2012). While much research has focused on the academic performance of community college students at four-year institutions, less attention has been paid to issues of student persistence (e.g., Townsend & Wilson, 2006; Glass & Harrington, 2010). At the same time, persistence and graduation may have a greater impact on students' long-term goals than their performance at an undergraduate institution. The purpose of this paper is to develop a model to predict community college transfer students' persistence at a four-year, online university. In particular, this paper operationalized persistence as students' re-enrollment in the immediate next semester after their first semester of transfer to the four-year institution. This measure of early persistence has been found to be a strong predictor of progress through the institution to graduation.

Theoretical Frame

Research on undergraduate persistence has been guided by two theoretical perspectives on student attrition. The first is Tinto's (1975; 1987) *Student Integration* Model, which identifies four aspects of student-institution interactions that affect persistence. Specifically, Tinto emphasized students' *background characteristics* and academic *goal commitments* and their effects on students' *academic* and *social integration* into the transfer institution as key to preventing student attrition. Background characteristics include students' demographic attributes, family backgrounds, and experiences prior to college (Tinto, 1975). Goal commitments include learners' motivations for degree pursuit and educational expectations as well as *institutional commitment* to a particular university. Academic and social integration are a consequence of students' interactions with a variety of institutional features over time (e.g., office hours, extracurricular activities). Tinto emphasizes the central importance of students' institutional integration, both academic and social, by saying, "we learned that involvement matters and that it matters most during the first critical year of college" (Tinto, 2006, p. 3; Upcraft, Gardner, & Barefoot, 2005).

At the same time, academic and social integration into a transfer institution is not a given for many students. Building on Tinto's earlier work (1975), Bean and Metzner (1985) developed a model of attrition, reflecting the experiences of non-traditional undergraduate students. Non-traditional undergraduate students are those who are older (i.e., 25 and above; Stewart & Rue, 1983), enrolled part-time, non-residential, commute to campus, or represent some combination of these characteristics (Bean & Metzner, 1985). Understandably, this population of students is thought to undergo a socialization process different from that of traditional students like those conceptualized in Tinto's model (1975). Non-traditional students may have different experiences with and potential for institutional commitment and academic and social integration. Bean and Metzner (1985) suggest that this variation may be, in part, because older students exhibit more characteristics associated with maturity. They may, therefore, be less open to socialization processes. Further, because non-traditional students spend more time off-campus they may have more limited contact with socializing agents (e.g., faculty, peers) (Chickering, 1974). While non-traditional students may be less interested in an institution's social culture, they may be more concerned with an institution's academic offerings and credentialing.

Juxtaposing the experiences of traditional and non-traditional learners, for non-traditional students there may be (a) more limited interactions with faculty and peers as well as with college services (i.e., more limited social integration, as per Tinto, 1975), (b) similarity in academic focus and experience (i.e., work experience that parallels the classroom), (c) much greater interaction with the external, non-institutional environment (e.g., work and family commitments), (Bean & Metzner, 1985), and (d) different expectations of the university and the classroom (Houser, 2005).

Based on the differences identified between traditional and non-traditional students, Bean and Metzner (1985) conceptualized students' decisions to drop out as predicated on four general types of variables. The first of these are *background factors*, including students' demographics, past academic performance, and educational goals and expectations. The second consideration is that of students' *academic performance,* encapsulating learners' grades, study habits, and pursuit of a major at the transfer institution. The third set of factors reflects students' *intent to leave*. Considered to be more psychological, these factors include students' goal commitment, perceived utility of a desired credential, and institutional satisfaction. Finally, unique to this model, is the inclusion of *external factors* that may have a direct effect on students' decisions to drop-out; these include finances, out-of-school work, and family commitments (Bean & Metzner, 1985).

The relation between the four factors identified may be said to be compensatory rather than independent. First, if students' academic outcomes are low, they may, nevertheless, persist, compensating for achievement deficits with high levels of psychological commitment. Further, even when academic performance is low, students will persist if external factors support their continued enrollment. Conversely, when external factors do not support persistence, for non-traditional students, even high academic performance may not be sufficient to ensure continued enrollment. As non-traditional students are much more closely affiliated with non-institutional environments than are traditional students residing on university campuses (Bean & Metzner, 1985; Metzner, 1984), Bean and Metzner suggest that—for non-traditional students—external factors have a much more pronounced effect on attrition decisions than do academic factors. At the institution level, understanding and supporting non-traditional student persistence may be particularly challenging as it may, in large part, be attributed to environmental factors that the institution may not be aware of or able to control.

Literature Review

In light of Bean and Metzner's model identifying factors that may challenge the academic persistence of non-traditional students, a number of studies have investigated the nature of transfer students' retention at four-year universities. Much of the research has focused on tabulating the attrition, persistence, and graduation rates of community college transfer students and comparing these to the performance of native students (Bean, 1980; Bean & Metzner, 1985).

A report from the National Center of Educational Statistics (NCES; Radford, Berkner, Wheeles, Shepherd, & White, 2010) examined students' persistence not only at their institution of first enrollment (i.e., community college) but also at their transfer destination. NCES determined that for students starting their post-secondary education in community college, six years later 12.9% were still enrolled at the less-than-four-year institution, 6.7% were still enrolled at a four-year institution, and 46.0% were no longer enrolled at any institution, having left without earning a credential. Based on a 2004 cohort, whose enrollment and graduation status were computed in 2009, 54% had succeeded in either earning a credential or were still enrolled in a four-year institution. This number can be compared to the 2004 native cohort of students starting at four-year institutions, only 23.6% of whom had dropped out prior to 2009 without having earned a credential. As such, the persistence and graduation rate of native students at a four-year institution (76.4%) has been found to be over 20 percentage points higher than that of community college transfer students.

Looking more specifically at year-to-year retention, while rates for native students are well documented, little is known about the year-to-year retention of transfer students at either the community college or the four-year institution. From the first to the second year, first time freshmen have been found to be retained at a rate of 76.3% according to Herzog (2005), a bit higher than findings by Bartlett and Abell (1995), who place first-to-second year student retention between 60 and 70 percent. At the same time, rates of retention have been found to vary widely across institutions (Hagerdon, 2005; Summerskill, 1962) and demographic categories (Nora, Barlow, & Crisp, 2005). Although generally, first-to-second year retention has been found to be lower at the community college level than at four-year institutions (e.g., SREB Fact Book, 2007), to our knowledge, no robust, national statistics on community college transfer students' first-to-second year retention exist.

Research on community college and nontraditional students is complicated by conflicting definitions of *retention* and *persistence* as well as by variable standards for calculating persistence. For example, retention rates for federally mandated reporting are computed for Fall-enrolled, first-time freshmen and thus may exclude many non-traditional students, students starting in a non-Fall cohort, or not working toward a degree. For students enrolled part-time, the window within which retention is calculated may not be sufficiently large to capture their progress toward a degree. From an institutional perspective, students who are not retained may be transferring to alternative institutions or stopping out (i.e., discontinuing their education for a variety of external reasons, beyond the institution's control). Shapiro et al. (2010) describe the challenges associated with computing retention as resulting from the diversity of pathways transfer students may pursue in obtaining a credential.

Predicting Retention

Rather than computing a specific rate of retention, the focus of this paper was on modeling or predicting retention. A limited number of studies have adopted such an approach.
Based on a comprehensive review of the persistence literature, Peltier, Laden, and Matranga (1999) determined gender, race/ethnicity, socioeconomic status, high school GPA, college GPA, and their interactions to be most strongly associated with persistence. Whereas findings about the role of gender in persistence have been mixed (Reason, 2009; St. John et al., 2001), race/ethnicity and prior academic achievement have been robust predictors of persistence (e.g., Astin, 1997; Tross, Harper, Osher, & Kneidinger, 2000; Levitz, Noel, & Richter, 1999).

Murtaugh, Burns and Schuster (1999) used survival analysis to examine traditional, first-time freshmen retention between 1991 and 1996. In their investigation, 25% to 35% of the cohort examined had interrupted enrollment within this six-year period. Specifically, 13.5% of students stopped out for a single term, 10.8% stopped out for two terms, and 1.8% stopped out for three terms, after which they were required to undergo a readmission process. Minority students had a higher rate of withdrawal than did white students; also associated with withdrawal were age, high school GPA, first quarter GPA, area of study, and participation in freshman orientation. In summarizing their work, Murtaugh et al. (1999) highlighted the importance of pre-college characteristics in predicting persistence.

Wetzel, O'Toole, and Peterson (1999) used logistic regression to predict a dichotomous outcome variable (i.e., retained or not retained). Retention was significantly predicted primarily by academic factors, including GPA and the ratio of credit hours students earned to those they attempted. This ratio, termed *course efficiency,* will be the focus of this investigation.

Predicting Retention for Nontraditional Students

While prior work has focused on predicting traditional student retention, more limited work has focused on the persistence of community college transfer students (e.g., Otero, Rivas, & Rivera, 2007). Based on theoretical work (Astin,1975; Bean & Metzner, 1985; Tinto, 1975), it is reasonable to expect that community college transfer students' persistence may be affected by factors different from those affecting traditional learners. First, given that much of the literature on community college transfer students has focused on students' academic preparedness (Carlan & Byzbe, 2000), learners' prior academic experiences may be particularly important to examine. For community college transfer students, prior academic experience would include not only work from high school, but also college-level course work completed at the community college. Second, community college transfer students may be considered to be nontraditional learners. As such, these students may have strong, external, nonuniversity connections (e.g., family, work commitments, Bean & Metzner, 1985). Given the importance of such external factors in predicting nontraditional student persistence, it may be particularly important to examine community college transfer students' background characteristics and how these relate to academic factors at the four-year institution in predicting retention.

Wang (2008) used logistic regression and found that the probability of graduating with a bachelor's degree for students starting at community college was predicted by gender, SES, high school curricula, educational expectations, community college GPA, college involvement, and math remediation. Persistence, prior to graduation, was predicted by community college GPA and locus of control. Similar to Wang (2008), this study will look to students' demographic characteristics and community college background factors (i.e., course-taking behaviors) to predict persistence or next-semester re-enrollment.

The target outcome in this study is first year re-enrollment. While this measure has been found to be associated with students' ultimate graduation, it may also be considered to be a metric of transfer students' *fit* or institutional integration in the first year of transfer. Measuring fit in the first year of transfer may be particularly important in understanding community college transfer student success and persistence. Krieg (2010) examined students at Western Washington University, an institution with a substantial population of community college transfer students comprising each class, and found that native (i.e., nontransfer) students were more likely to graduate, even after controlling for demographic characteristics and prior academic performance. Krieg (2010) compares the experience of community college students to that of freshmen at a four-year university. Community college transfer students may have difficulties adjusting to a new learning context; this may result in early attrition if students consider themselves to be incompatible with their new, four-year university environment.

Community college transfer student fit has most commonly been examined by comparing students' academic performance at the community college vis-à-vis the four-year university. Typically, a decrease in performance (i.e., GPA) after transfer has been identified and termed transfer shock (Cejda, Kaylor, & Rewey, 199; Townsend, McNemy, & Arnold, 1993). However, Krieg (2010) has suggested that students' feelings of misfit at the four-year university may manifest more profoundly not in GPA decline, but as rapid attrition from the four-year institution. In other words, community college students who have difficulties integrating into their transfer institution may withdraw quickly, as early as their first semester or first year of transfer. Indeed, the majority of transfer student attrition has been found to occur in the first year of transfer, when students are new to the university setting (Ishitani, 2008; Laanan, 2001; Monroe, 2001).

Krieg (2010) further cautions that early attrition does not occur only among low performing students. Even high performing community college transfer students are more likely to drop out than are their native counterparts. This may be because transfer students have less immediate affiliation and integration into the transfer institution or because community college students may experience unique academic challenges upon transfer. For instance, transfer students are typically required to take a number of prerequisite courses before entering into a major (Krieg, 2010). Such findings are broadly consistent with Bean and Metzner's (1985) proposition that for nontraditional learners, even academic success may not be sufficient to ensure persistence.

Present Study

The purpose of the present study was to examine a predictive model of community college transfer students' first-year persistence at a four-year, online university. In this study, persistence is defined as students' re-enrollment in the immediate next semester after their first semester of transfer. Re-enrollment, as a measure of early persistence, may be considered to reflect both students' integration into the transfer institution and commitment to continued enrollment until graduation.

Predictors of Community College Transfer Student Re-Enrollment

In this study, three types of factors were examined in predicting re-enrollment. These were (a) *demographic factors*, (b) *community college background factors,* including course- taking behaviors, and (c) *course efficiency*, a novel variable developed as a summative measure of transfer students' progress at the community college. Demographic factors, found to impact persistence in prior research, were examined. These included age, gender, and race/ethnicity (see Reason, 2009, for a review). Community college background factors were examined to capture learners' prior academic experience. Rather than examining GPA, a summative measure of academic performance, well-established as predictive of persistence in prior research (e.g., Reason, 2009; Wang, 2008), students' course-taking at the community college was examined for this research study. Examining course-taking allowed for a more fine-grained analysis of community college transfer students' academic preparedness. For example, math performance at the community college and the need for remediation have been found to be particularly strong determinants of students' success and persistence upon transfer (e.g., Wang, 2008). For transfer students, community college often represents a first encounter with college-level course work. As such, we considered an in-depth examination of course-taking behaviors at the community college and their association with persistence upon transfer.

Finally, a novel metric, *course efficiency*, was introduced as a summative measure of students' community college backgrounds. Students' course efficiency reflects the ratio of credits earned at the community college to the number of credits attempted. Although used infrequently in prior research (e.g., Wetzel et al., 1999), course efficiency captures students' track records in making successful progress toward meeting self-determined academic milestones. Rather than focusing on performance, course efficiency contextualizes students' course completion in terms of their own goals and expectations for progress.

For nontraditional students, course efficiency, or effectiveness and persistence in pursuing academic goals, may better reflect academic success than does absolute achievement, measured by GPA. While GPA only considers academic performance, course efficiency may better capture the real-world consequences of students' deviation from planned academic pathways in terms of time and added financial cost. Additional time and cost may pose particular threats to the successful degree completion of nontraditional students. For nontraditional students, decisions to persist have been conceptualized as a weighing of the costs and benefits of attending college (Tinto, 1986; Braxton & Hirschy, 2000). For instance, students may weigh the costs—financial, psychological, and social—of stepping out of the workforce to pursue a degree vis-à-vis the long-term earning potential associated with earning a credential. Course efficiency can be thought to reflect this type of cost-benefit calculus by presenting a single metric of the relative expediency of students' academic pathways judged against their self-determined goals. In the present study, we were interested in examining the extent to which course efficiency at the community college may predict re-enrollment at the transfer institution over and above the respective roles of demographic and community college background factors.

Online Context

In this study, predictors of persistence were examined in a unique context: an online university. There are a number of reasons why distinct models of persistence need to be developed for students learning online. For one, even compared to nontraditional students, these learners may struggle with institutional integration because they may have more limited direct, in-person contact with socializing agents like faculty and peers. For another, their institutional interactions may be different in nature (e.g., asynchronous and requiring a greater degree of written communication). Online learners' institutional integration may further be hampered by difficulties connecting with an institution with limited physical infrastructure and facilities.

Learner factors likewise contribute to difficulties with online institutional integration. First, students enrolling in an online university often do so because of limited resources. Specifically, students may pursue online learning because of work or family commitments that prevent their enrollment in a brick-and-mortar institution. In turn, such external commitments may interfere with students' institutional integration and make the cost of pursuing an education particularly high. Further, when learning online, it may be more difficult for students to quickly and directly receive help from professors or peers. Indeed, learning online has been found to require students to be more responsible and self-directed in their learning (Diaz & Cartnal, 1999; Vonderwell, 2003); however, this may also serve to limit students' openness to institutional integration.

Taken together, both institutional and student factors associated with online learning may combine to reduce students' likelihood of integration into an online university, contributing to attrition. Risk of withdrawal may increase all the more for students transferring from a brick-and-mortar community college to a four-year, online university. Both going from a community college to a four-year institution and transitioning from an in-person learning environment to learning online may contribute to students perceiving a mismatch between themselves and their transfer destination, resulting in withdrawal. Put simply, students unfamiliar with the demands of online learning may make a decision that it is "not for them," and therefore fail to persist.

Considering the increased risk of withdrawal associated with learning online, particularly for community college transfer students, examining persistence in this population is of particular importance. In this study, students' demographic factors, community college course-taking behaviors, and course efficiency were used to predict re-enrollment among community college transfer students enrolled at a four-year university.

This research study addresses the following research questions:

1. To what extent do demographic characteristics predict re-enrollment for community college transfer students at an online, four-year university?
2. Controlling for demographic factors, to what extent do community college course-taking behaviors predict re-enrollment for community college transfer students at an online, four-year university?
3. To what extent does course efficiency at the community college predict re-enrollment for community college transfer students at four-year, online university, once demographic and community college background factors are controlled for?

Methods

Participants

Participants were 8,200 community college transfer students enrolled in a four-year online university during an 11-semester window (Fall 2005 – Spring 2011). Students had transferred from two area community colleges, constituting the two largest institutions of origin for students transferring to the four-year university. The sample was majority female (57.5%, n=4715, 41.3% male, n=3387) and had a mean age of 28.68 (*SD*=8.43) years in their first semester of transfer. The sample was racially and ethnically diverse, 21% White (n=1727), 43.2% African American (n=3546), 10.2% Hispanic/Latino (n=837), 10.6% Asian (n=866), 0.90% Native American (n=76), with 14.0% of the sample not specifying a race.

Measures

Data for this study were collected through a cross-institutional collaboration between two partner community colleges and a four-year university serving as the transfer destination. Student records were matched across institutions and a combined dataset, reflecting students' records prior to and following transfer, was developed. Demographic factors, courses taken at the community college, community college course efficiency, and first-term GPA at the transfer institution were all used to predict students' persistence, measured as next semester re-enrollment.

Independent variables. Four types of predictor variables were used in analyses: learners' (a) *demographics*, (b) community college *course-taking behaviors*, (c) *course efficiency*, and (d) *first-term GPA* at the transfer institution. Each predictor variable will be explained below.

Learner characteristics. Four primary demographic variables were considered. These were (a) *age*, (b) *gender*, (c) *race/ethnicity*, and (d) *marital status*.

Course-taking. A variety of course-taking behaviors at the community college were examined in the model. These included the subject areas of courses taken and course attributes (e.g., honors).

Subject areas. Whether or not students took courses in four different subject areas at the community college was examined. Specifically, dichotomously coded variables for whether or not students had taken courses in *math, English, speech,* or *computers* were entered as predictors into the model. Courses in these four subject areas were considered to be gateways, mandatory for students seeking a credential and granting students access to higher-level classes across subject areas.

Course attributes. Based on an examination of community college student records, three course attributes distinguishing the academic offerings students experienced in community college were considered. Specifically, these were whether or not students had taken a course designated as (a) honors or not or (b) developmental or not at the community college. Also, a dichotomous variable for whether or not students had ever taken an (c) online course was entered into the model. This was considered to be a particularly interesting predictor to examine as our sample included students transferring from brick-and-mortar community colleges to an online institution. Having prior experience in online learning may have improved students' persistence by increasing their familiarity with learning online.

Course efficiency. Course efficiency, defined as the ratio of credits earned at the community college to credits attempted, was included as a summative measure of students' community college experiences. In modeling students' community college backgrounds, we considered it important to reflect how many credits students had earned at the community college relative to what their aims were. Credits earned was considered to be an important metric, having implications for timely graduation, and allowing the experiences of students enrolled in community college for varying durations, prior to transfer, to be directly compared.

First-term GPA. Finally, as per prior research, first-term GPA at the transfer institution was entered into the model as a predictor of re-enrollment.

Dependent variable. Individual difference factors, community college course-taking behaviors, course efficiency, and first-term GPA at the transfer institution were collectively used to predict students' re-enrollment. Re-enrollment referred to students' enrollment in the subsequent semester, following their first semester of transfer. Logistic regression was used to predict re-enrollment. Re-enrollment was modeled based on variables associated with students' academic careers at both the community college and the four-year university.

Results

Three hierarchical logistic regressions were run predicting the dichotomous outcome variable, re-enrollment. The first considered only students' demographic information and first-term GPA at the transfer institution as predictive of re-enrollment. The second regression examined the extent to which demographic information, community college course-taking, and first-term GPA at the transfer institution were predictive of re-enrollment. Finally, a full model of student re-enrollment was examined, including demographic factors, community college course-taking, course efficiency, and first-term GPA as predictors.

Demographic Characteristics

In the first model, students' demographic characteristics (i.e., gender, age at transfer, race/ethnicity, and marital status) were entered at Step 1 and first-term GPA at the transfer institution were entered at Step 2. The model was overall significant, $\chi^2(9) = 655.51$, $p<.001$. Specifically, 71.2% of participants were correctly classified as having re-enrolled or not. Looking at pseudo-R^2 measures, between 7.77 and 10.81% of variance in re-enrollment was explained by the model, according to Cox and Snell's R^2 and Nagelkerke's R^2, respectively.

Looking at the individual predictors in the model, gender (β = 0.16, *SE*(β) = 0.05, $p<0.01$), age (β = -0.01, *SE*(β) = 0.00, $p<0.001$), marital status (β = 0.31, *SE*(β) = 0.07, $p<0.001$), and Black (β = 0.29, *SE*(β) = 0.06, $p<0.001$) and Asian ethnicity (β = 0.19, *SE*(β) = 0.09, $p<0.05$) were all significantly associated with re-enrollment, as was first-term GPA (β = 0.44, *SE*(β) = 0.02, $p<0.001$).

Community College Course-taking Behaviors

A second model was run to examine the extent to which community college course-taking behaviors could predict re-enrollment over and above students' demographic characteristics. Hierarchical regression was used, with demographic characteristics entered at Step 1, community college course-taking behaviors entered at Step 2, and first-term GPA at the transfer institution entered at the final step. The model was overall significant, X^2 (16) =695.95, $p<.001$, with 71.38% of students correctly classified as re-enrolled or not. The variance explained was between 8.23% according to Cox and Snell R^2 and 11.45% according to Nagelkerke's R^2.

In terms of the individual predictors in the model, while demographic characteristics, particularly gender (β = 0.15, *SE*(β) = 0.05, $p<0.01$), age at first transfer (β = -0.01, *SE*(β) = 0.00, $p<0.01$), marital status (β = 0.30, *SE*(β) = 0.07, $p<0.001$), and identification as African American (β = 0.26, *SE*(β) = 0.07, $p<0.001$) were significant predictors of re-enrollment, course-taking in specific subject areas was not (English, $p=0.59$; Speech, $p=0.38$; Computers, $p=0.38$). However, enrollment in math courses was significantly associated with re-enrollment (β = 0.13, *SE*(β) = 0.06, $p<0.05$). While having taken honors courses was not a significant predictor in the model ($p=0.29$), having taken developmental courses (β = 0.15, *SE*(β) = 0.06, $p<0.01$) and classes online (β = 0.12, *SE*(β) = 0.05, $p<0.05$) were both significantly associated with re-enrollment. Finally, first-term GPA was a significant predictor in the model (β = 0.44, *SE*(β) = 0.02, $p<0.001$).

Course Efficiency

The final model examined the full cadre of predictors, including students' demographic characteristics, course-taking behaviors at the community college, community college course efficiency, and first-term GPA. The model was overall significant, with 71.31% of students successfully classified as re-enrolling or not. According to Cox and Snell's R^2 8.35% of variance in re-enrollment was explained, whereas according to Nagelkerke's R^2 11.61% of variance was explained.

Looking to the individual predictors, demographic characteristics [i.e., gender, β = 0.15, *SE*(β) = 0.05, $p<0.01$; age, β = -0.01, *SE*(β) = 0.00, $p<0.05$; marital status, β = 0.30, *SE*(β) = 0.07, $p<0.001$; and identification as African American, as compared to White, β = 0.25, *SE*(β) = 0.07, $p<0.001$] were all significant predictors in the model. Further, while course-taking in particular subject areas was not significantly predictive of re-enrollment [i.e., English, $p=0.72$; Math, $p=0.05$; speech, $p=0.20$; computers, $p=0.41$], enrolling in developmental education, β = 0.14, *SE*(β) = 0.06, $p<0.05$, and taking courses online, β = 0.10, *SE*(β) = 0.05, $p<0.05$, were significant predictors in the model. Finally, both course efficiency,

β = -0.00, $SE(\beta)$ = 0.00, p<0.01, and first-term GPA at the transfer institution, β = 0.46, $SE(\beta)$ = 0.02, p<0.001, were significant predictors of re-enrollment.

Interestingly, minority status (i.e., reporting African American race/ethnicity) in the full model as well as in models prior was positively associated with retention. As compared to White students, African American students were 1.25 times more likely to re-enroll, holding all other factors consistent in the model. This is a notable finding given that minority status has often been considered to be a risk factor for college completion (Murtaugh et al., 1999). Further, course efficiency in the final model was a slightly negative predictor of re-enrollment at a four-year institution.

Discussion

This study examined factors associated with community college transfer student persistence at a four-year online university. Variables associated with performance at both the community college and the institution of transfer were examined. Consistent with prior research, demographic variables were found to be significantly associated with re-enrollment. In particular, both age and marital status were found to have a positive association with persistence. While it may be the case that nontraditional students, who are older and face the demands of balancing career and family, may struggle with prioritizing academics, it seems that for these students, added maturity serves to improve academic achievement. Older students, potentially working while in school, may also benefit from their work paralleling activities in the classroom.

An important finding in this study is that across models, African American students were more likely to persist than were their White counterparts. While prior work has associated minority status with lower academic achievement (Brown, Brown, Beale, & Gould, 2014; Gentry, 2014; Farmer & Hope, 2015), it may be the case that minority students are motivated to persist, even in the face of academic challenge. Alternately, it may be the case that when provided with the autonomy and flexibility afforded by online learning, minority and other at-risk students are able to successfully persist in their academic pursuits. Further examining the interaction between measures of both achievement and persistence for at-risk students is a promising avenue for further investigation.

Course-taking at the community college was found to have a limited effect on persistence. Specifically, students' course-taking in particular subject areas was not a significant predictor of re-enrollment. This may have been because all students were required to take core courses in specific subject areas, limiting their variability. Alternately, it may have been the case that not taking courses in particular subject areas was associated with students' intent to transfer and to complete required courses at the transfer institution. Further work should consider how constellations of courses taken or majors pursued may impact persistence at the transfer institution. Indeed, transfer students' academic performance has been found to differ across subject areas (Cejda, Kaylor, & Rewey, 1998), however the associations between areas of study and persistence have not been as well documented.

At the same time, students' enrollment in a developmental course and an online course during their time in community college was positively associated with re-enrollment at the transfer institution. While taking a developmental course may be associated with limitations in community college students' preparedness, this may also signal students' commitment to building fundamental skills necessary for academic success. Online course-taking at the community college may have served to prepare students for the demands of learning in a fully-online university. Further, taking courses online may have helped students decide whether learning online was a good fit for their needs and desires as a student. Both completing developmental course work and attempting an online course at the community college may be indicative of students' forethought and planning to transfer to a four-year institution.

Course efficiency at the community college was a strong predictor of persistence at the transfer institution. In addition to reflecting students' abilities and motivation to complete course work successfully, course efficiency may have provided transfer students with an advantage upon transfer. Specifically, students with a high rate of course efficiency may have completed more credits at the community college in a shorter amount of time. This means that they may have transferred to a four-year institution with more time and energy ready to dedicate to degree pursuit—a particular concern for non-traditional learners. More needs to be known about whether course efficiency at the community college best reflects students' academic abilities, skills in successfully completing course work, or some combination of these two factors. However, considering the strength of course efficiency in predicting early persistence, more work is needed to better understand this factor and to provide community college students with interventions to support efficient and efficacious course-taking.

The strongest predictor of persistence was first-term GPA at the transfer institution. This was an expected finding. The strong association between first-term GPA and persistence may come, in part, from these measures being captured at the same institution, within a close period of time. Further, these two measures may be closely associated because transfer students consider first-term GPA in their decisions to persist or withdraw. Students may use first-term GPA as a barometer not only of their academic performance but of their institutional fit as well (Krieg, 2010). Results in this study are consistent with prior research that has considered first- term GPA to be a measure of students' institutional integration.

Conclusions and Implications

This study contributes to the literature on community college transfer student success in at least four ways. First, it tracks the academic trajectories of a robust student sample across institutions, from community college to a four-year university. Second, this study examines student persistence in a unique academic context—in a four-year, online university. Third, in addition to using well-documented predictors of persistence, including demographic factors and first-term GPA, this study examined the role of a novel factor, course efficiency, in predicting student persistence. Finally, by examining specific courses taken at the community college, this study provides insight into potential areas for intervention to improve community college students' transition from a two-year to a four-year institution.

Authors Note

This research was supported by the University of Maryland University College, Montgomery College, and Prince George's Community College as well as the Kresge Foundation.

In addition, the authors wish to thank Drs. Peter Shea and Ben Arbaugh for their support.

Correspondence concerning this article should be addressed to Institutional Research, University of Maryland University College, 3501 University Boulevard East, Adelphi, MD 20783.
Contact: denise.nadasen@umuc.edu

References

Astin, A. W. (1997). The changing American college student: Thirty-year trends, 1966-1996. *The Review of Higher Education*, *21*(2), 115–135.

Bean, J. P., & Metzner, B. S. (1985). A conceptual model of nontraditional undergraduate student attrition. *Review of Educational Research*, *55*(4), 485–540. http://doi.org/10.2307/1170245

Bartlett, C., & Abell, P. (1995). Understanding the transfer student: Or are we? Paper presented at the Annual Meeting of the American Association of Collegiate Registrars and Admissions Officers. Tucson, AZ.

Brown, C. S., Brown III, U. J., Beale, R. L., & Gould, J. K. (2014). Factors influencing academic success and self-esteem among diverse college students. *International Journal of Education Research*,*9*(1). 1-14.

Cabrera, A. F., Nora, A., & Castaneda, M. B. (1993). College persistence: Structural equations modeling test of an integrated model of student retention. *Journal of Higher Education*, 123-139.

Cabrera, A. F., Castaneda, M. B., Nora, A., & Hengstler, D. (1992). The convergence between two theories of college persistence. *The Journal of Higher Education*, 143-164.

Carlan, P. E., & Byxbe, F. R. (2000). Community colleges under the microscope: An analysis of performance predictors for native and transfer students. *Community College Review*, *28*(2), 27-42.

Cejda, B. D., Kaylor, A. J., & Rewey, K. L. (1998). Transfer shock in an academic discipline: The relationship between students' majors and their academic performance. *Community College Review*, *26*(3), 1-13.

Diaz, D. P., & Cartnal, R. B. (1999). Students' learning styles in two classes: Online distance learning and equivalent on-campus. *College Teaching*, *47*(4), 130-135.

Farmer, E. D., & Hope, W. C. (2015). Factors That Influence African American Male Retention and Graduation the Case of Gateway University, a Historically Black College and University. *Journal of College Student Retention: Research, Theory & Practice*, *17*(1), 2-17.

Gentry, R. (2014). Sustaining college students' persistence and achievement through exemplary instructional strategies. *Research in Higher Education*. *24*, 1-14.

Glass, C. G., & Harrington, A. R. (2002). Academic performance of community college transfer students and "native" students at a large state university. *Community College Journal of Research and Practice*, *26*(5), 415–430. http://doi.org/10.1080/02776770290041774

Hagerdon, L.S. (2005) How to define retention: A new look at an old problem. In A. Seidman (Ed.) *College Student Retention: Formula for Student Success*. Greenwood Publishing Group: New York.

Herzog, S. (2005). Measuring determinants of student return vs. dropout/stopout vs. transfer: A first-to-second year analysis of new freshmen. *Research in Higher Education*, *46*(8), 883–928. http://doi.org/10.1007/s11162-005-6933-7

Houser, M. L. (2005). Are we violating their expectations? Instructor communication expectations of traditional and nontraditional students. *Communication Quarterly, 53(*2), 213-228.

Hu, S., & John, E. P. S. (2001). Student persistence in a public higher education system: Understanding racial and ethnic differences. *The Journal of Higher Education*, *72*(3), 265–286. http://doi.org/10.2307/2649332

Ishitani, T. T. (2008). How do transfers survive after "transfer shock"? A longitudinal study of transfer student departure at a four-year institution. *Research in Higher Education*, *49*(5), 403-419.

Krieg, J. M. (2010). Academic performance of native and transfer students. *Office of Survey Researcher*, Paper 12, Western Washington University. Retrieved from: http://www.wwu.edu/osr/documents/OSRReport2010-02.pdf

Laanan, F. S. (2001). Transfer student adjustment. *New Directions for Community Colleges*, *2001*(114), 5-13.

Levitz, R. S., Noel, L., & Richter, B. J. (1999). Strategic moves for retention success. *New Directions for Higher Education*, (108), 31.

Marks, J. L. (2007). *SREB Fact Book on Higher Education 2007*. Southern Region Educational Board: Atlanta, GA.

Monroe, A. (2006). Non-traditional transfer student attrition. *The Community College Enterprise*, *12*(2), 33-47.

Murtaugh, P. A., Burns, L. D., & Schuster, J. (1999). Predicting the retention of university students. *Research in higher education*, *40*(3), 355-371.

Nora, A., Barlow, L., & Crisp, G. (2005). Student persistence and degree attainment beyond the first year in college. *College student retention: Formula for success*, 129-153.

Otero, R., Rivas, O., & Rivera, R. (2007). Predicting persistence of Hispanic students in their 1st year of college. *Journal of Hispanic Higher Education*, *6*(2), 163–173. http://doi.org/10.1177/1538192706298993

Pascarella, E. T., & Terenzini, P. T. (1980). Predicting freshman persistence and voluntary dropout decisions from a theoretical model. *The Journal of Higher Education*, *51*(1), 60–75. http://doi.org/10.2307/1981125

Peltier, G. L., Laden, R., & Matranga, M. (1999). Do high school athletes succeed in college: A review of research. *The High School Journal*, *82*(4), 234–238.

Radford, A. W., Berkner, L., Wheeless, S. C., & Shepherd, B. (2010). *Persistence and Attainment of 2003-04 Beginning Postsecondary Students: After 6 Years*. National Center for Education Statistics: Washington, DC.

Reason, R. D. (2009). An examination of persistence research through the lens of a comprehensive conceptual framework. *Journal of College Student Development*, *50*(6), 659–682. http://doi.org/10.1353/csd.0.0098

Seidman, A. (2005). *College Student Retention: Formula for Student Success*. Greenwood Publishing Group.

Shapiro, D., Dundar, A., Chen, J., Ziskin, M., Park, E., Torres, V., & Chiang, Y. C. (2012). *Completing College: A National View of Student Attainment Rates.* National Student Clearinghouse: Washington, DC.

Spady, W. G. (1970). Dropouts from higher education: An interdisciplinary review and synthesis. *Interchange*, *1*(1), 64–85. http://doi.org/10.1007/BF02214313

Stewart, S. S., & Rue, P. (1983). Commuter students: Definition and distribution. *New Directions for Student Services*, *1983* (24), 3–8. http://doi.org/10.1002/ss.37119832403

Summerskill, J. (1962). Dropouts from college. In S. Nevitt (Ed), *The American college: A psychological and social interpretation of the higher learning* (pp. 627–657). John Wiley & Sons: Hoboken, NJ.

Tinto, V. (1975). Dropout from higher education: a theoretical synthesis of recent research. *Review of Educational Research*, *45*(1), 89–125. http://doi.org/10.2307/1170024

Tinto, V. (1987). *Leaving College: Rethinking the Causes and Cures of Student Attrition.* University of Chicago Press: Chicago, IL.

Tinto, V. (2006). Research and practice of student retention: What next? *Journal of College Student Retention: Research, Theory & Practice*, *8*(1), 1–19.

Tross, S.A., Harper, J.P., Osherr, L.W. & Kneidinger, L.M. (2000). Not just the usual cast of characteristics: using personality to predict college performance and retention. *Journal of College Student Development, 41*(3), 325-336.

Townsend, B. K., & Wilson, K. (2006). “A hand hold for a little bit”: Factors facilitating the success of community college transfer students to a large research university. *Journal of College Student Development*, *47*(4), 439–456. http://doi.org/10.1353/csd.2006.0052

Upcraft, M., Gardner, J., & Barefoot, B. (2005). *Challenge and support: Creating climates for first-year student success*. Jossey-Bass: San Francisco.

Vonderwell, S. (2003). An examination of asynchronous communication experiences and perspectives of students in an online course: A case study. *The Internet and Higher Education*, *6*(1), 77-90.

Wang, X. (2012). Academic performance of community college transfers: psychological, sociodemographic, and educational correlates. *Community College Journal of Research and Practice*, *36*(11), 872-883.

Exploring the Relationships between Facilitation Methods, Students' Sense of Community, and Their Online Behaviors

Krystle Phirangee, Carrie Demmans Epp, and Jim Hewitt
University of Toronto

Abstract

The popularity of online learning has boomed over the last few years, pushing instructors to consider the best ways to design their courses to support student learning needs and participation. Prior research suggests the need for instructor facilitation to provide this guidance and support, whereas other studies have suggested peer facilitation would be better because students might feel more comfortable learning and challenging each other. Our research compared these two facilitation methods and discovered that students participated more in instructor-facilitated online courses where they wrote more notes, edited and reread notes more, and created more connections to other notes than students in peer-facilitated online courses. We identified student activity patterns and described differences in how those patterns manifest themselves based on the facilitation method that was used. Our findings also show that instructor-facilitated online courses had a stronger sense of community than peer-facilitated online courses.

Introduction

Online enrollment has risen at the postsecondary level (Allen & Seaman, 2013; Gray, 2013) with courses being offered through a variety of online learning environments (OLE). These OLEs have been praised because they allow learners to work at a pace that is comfortable for them and supportive to their individual learning (Brooks, Demmans Epp, Logan, & Greer, 2011; Bolliger & Inan, 2012). Some learners also seem to favor online courses because they provide discussion forums where students can share resources, engage in discourse, and reflect on their ideas (Hewitt, 2005). These discussion forums allow students to have greater access to others' ideas as well as provide a space for all students to participate and share simultaneously in many discussions at their own pace, which gives shy or quiet students more time to think before contributing to discussions (Hewitt, 2005; Dzubinski, 2014). These advantages have contributed to the recent growth in distance education.

This rapid growth and the varied manner in which OLEs and their associated discussion forums can be used to support student learning means that we do not fully understand how student activities and discussion forum design influence one another. Moreover, the historical absence of analytics that describe student activities within OLEs or their relationship to theoretically relevant constructs, such as the learner's sense of community, limits our understanding of the learning processes and experiences of those who have enrolled in online courses. This limited understanding only serves to make the problems that are associated with OLEs increasingly pressing. These problems include high dropout rates (Carr, 2000; Kizilcec, Piech, & Schneider, 2013) that have been attributed to feelings of isolation and alienation among students due to their physical separation (Rovai, 2002a), a lack of interaction between students (Carr, 2000), miscommunication between students (Rovai & Wighting, 2005; Zembylas, 2008), and the clash of societal beliefs with students' personal and cultural beliefs (Rovai & Wighting, 2005). There is also a "need for an education system that helps people to help each other" (Ferguson & Bukingham Shum, 2012, p.8), since the isolation and disconnectedness of students in an OLE is the main reason for student attrition in online courses (Angelino, Williams, & Natvig, 2007).

To provide a more comfortable, safe, and positive online learning experience, instructors are encouraged to focus on the "the social nature of learning," which emphasizes the need for interactions and discussions between students (Hew, 2015, p.2). This fits with a social constructivist perspective, where learning is a process in which learners are able to construct new meanings through interaction and active involvement (Vygotsky, 1980). It emphasizes that collaboration and interaction with others produces deep and meaningful learning. This framework also emphasizes that students need to be a part of their own learning by recognizing that new knowledge and understandings are created when their own beliefs interact with those of others (Richardson, 2003), making the purpose of education enabling learners to work together to construct knowledge (Shackelford & Maxwell, 2012). Doing so allows learners to value their own learning and their peers' learning because interacting with various knowledge sets leads to new understandings. These types of student interactions can be supported through online discussion forums, which can be facilitated by the instructor or students. In these spaces, novices and experts work together, and everyone mentors and is mentored because everyone is a learner no matter their experience or expertise (Ferguson & Bukingham Shum, 2012).

Some scholars identify discussion forums as the place where a class-wide learning community develops (Arend, 2009), because it is here that students interact with the content, the instructor, and their peers to reshape and develop their knowledge (Song & McNary, 2011; Swan, 2009). Discussion forum interaction is "a necessary and fundamental process for knowledge acquisition and cognitive development" (Barker as cited in Song & McNary, 2011, p. 1). Students have also stressed that using discussion forums benefits their learning (Ertmer et al., 2007). Given this landscape, instructors are trying to foster a *sense of community* in their online courses, which refers to a feeling of belonging and interactivity among learners in an OLE (Rovai, 2002a; Liu et al., 2007; Ouzts, 2006). Fostering community can reduce feelings of disconnection and isolation, because being part of a community allows students to build camaraderie and engage in social reinforcement (Conrad, 2005; Gallagher-LePak, Reilly, & Killion, 2009), with interaction being critical to building a class-wide learning community online (Arend, 2009; Song & McNary, 2011; Swan, 2009) because it is thought to lead to deeper thinking (Hulon 2013; Larson & Kieper, 2002) and better student outcomes (Liu et al., 2007). Rovai (2002a) suggests that fostering community in an online course has the potential to minimize feelings of isolation, alienation, and disconnection online learners may experience.

However, there is a lack of consensus as to which strategies (such as instructor or peer facilitation) most effectively foster a sense of community in OLEs. This lack of consensus may be partly due to a lack of reliable analytics that detail how different strategies influence learner experiences. The development of these analytics could be used to address criticism about the lack of clear direction or empirically derived research that illustrates how to develop effective online communities (Liu et al., 2007). One of the areas that could be better informed by the use of analytics is how different facilitation methods influence student activities and support the development of a sense of community in an OLE, since there seem to be conflicting recommendations. Some work urges instructors to adopt the role of online facilitator: This involves clarifying course topics, keeping the discussions on track, introducing opposing views to students, helping students navigate the online platform, and emphasizing good online behavior (Hew, 2015). Others question whether the instructor should facilitate online discussions, because they feel it may be too time-consuming to oversee discussions properly and may unintentionally develop discussions that center on the instructor's comments (Correia & Baran, 2010; Light, Nesbitt, Light, & White, 2000). As a result, peer facilitation—in which students collaboratively control the discussions in an OLE (Bull, Greer, McCalla, & Kettel, 2001; Hew, 2015)—was proposed because students may increase their cognitive engagement when they recognize that their instructor is less engaged (Belcher, Hall, Kelley, & Pressey, 2014), and they may feel more comfortable asking for help, discussing their experiences, challenging and negotiating ideas, and sharing their views in a peer-facilitated online discussion rather than an instructor-led one (Bull, Greer, McCalla, & Kettel, 2001). According to Clarke and Bartholomew (2014), research on the role of the instructor has been inconsistent; they argue that discussions should be moderated, but how much moderation is needed from the instructor and the extent to which the instructor's participation matters is unclear.

While both peer and instructor facilitation have been argued for (Rourke & Anderson, 2002; Correia & Davis, 2007; Mazzolini & Maddison 2007; Arend, 2009; Hew, 2015), it is important to understand how these methods influence the creation of a sense of community and support students' learning needs in OLEs, especially with peer facilitation being minimally explored despite its popularity and potential (Baran & Correia, 2009; Clarke & Bartholomew, 2014). Our research aimed to contribute to the discussion by exploring the impact of peer and instructor facilitation on students' online behavior through a variety of analytics that include students' discussion forum posting patterns. Specifically, the following research questions are explored: (1) Which facilitation method (i.e., instructor or peer) helps to develop a stronger sense of community online? (2) How are students' online activity patterns related to instructor and peer facilitation methods?

An awareness of students' online activity patterns in weak and strong online communities and the facilitation method related to these communities is important in identifying the method that best fosters an online sense of community and that method's influence on students' online behaviors. Having this understanding along with the development of analytics that describe students' online behaviors holds the potential of offering directions and strategies to design online courses with the aim of providing more positive online learning experiences for students.

Method

As recommended by several scholars (Gunawardena, Lowe, & Anderson, 1997; Hiltz & Arbaugh, 2003), this study adopted a mixed-methods sequential-explanatory design (Creswell & Plano Clark, 2011). For the quantitative phase, we observed students' online activity patterns and asked those students to complete a questionnaire. For the qualitative phase we conducted semistructured interviews with students from the online courses.

Participants and Data Collection

This research took place at a public research university in a large metropolitan city in Canada. The data was collected from graduate students in the field of education who have taken online courses through the Pepper OLE. Pepper is a web-based collaborative workspace used by approximately 2,000 postgraduate education students every academic year. We gathered data and recruited participants through Pepper because it offers a variety of specialized knowledge building features and social networking tools to support the sharing of information to develop ideas.

Rovai’s (2002b) Classroom Community Scale (CCS) was randomly distributed to eight online graduate courses that used Pepper. These eight courses were then stratified by course format (i.e., seminar, independent study, and lecture) and instructor (i.e., different online instructors). Once stratified, the online courses with the same course format and different instructors were used in this study. Courses that did not meet this criterion were cut to ensure similar courses were being compared. For this study, six courses were selected for inclusion (N = 110), with each course having a different instructor; 47 students were spread across the three instructor-facilitated online courses, and 57 students were spread across the three peer-facilitated online courses. It is important to highlight that instructors from the instructor-facilitated online courses encouraged their students to be respectful to their peers and provide thoughtful online notes, but did not have any mandatory guidelines for their students with regard to participating in online discussions, whereas instructors from peer-facilitated courses had mandatory note guidelines for their students that included evaluation rubrics for note content, note length, and number of note contributions per week.

The CCS was used to determine which courses had a high or low sense of community. Once the high and low scoring communities were identified, we explored students’ online activity patterns by analyzing the logs that Pepper automatically collects. These logs store time-stamped records of online events (e.g., the creation of a note or when the Like button is pressed). Statistical data was pulled from the Pepper logs to understand students’ online behaviors and activity patterns. For the qualitative phase, semistructured interviews were conducted with graduate students from the participating online courses. Interview data in the qualitative phase focused on students’ views, experiences, and behavior in their online course.

Analysis and Validity

The CCS data was first cleaned to remove incomplete or duplicate responses. Then it was coded following the scale guidelines, and the highly reliable (α = .906) CCS score was calculated. The average response rate (M = 53.30%, SD = 15.60) was also calculated. Interestingly, the research community has yet to agree upon an acceptable minimum response rate, but it is not uncommon to see response rates below 20% (Fowler, 2009), with a response rate of 33% being typical (Nully, 2008). These statistics and the fact that nonresponse error does not have a strong relationship with survey error (Fowler, 2009) indicate that the response rate for the CCS was sufficient. Inferential statistics were then used to determine whether students’ perceptions of the sense of community that they felt within their online course differed.

The Pepper log files were analyzed from two perspectives. The first considered students’ activity levels (i.e., their raw activity counts for the term), and the second considered students’ system usage patterns (i.e., the ebb and flow in their activity levels from week to week). The analysis of student activity levels used inferential statistics to determine differences in the levels of various student activities. Comparisons were made based on students’ course membership and the facilitation method (i.e., instructor or peer) used within their course. Since Pepper tracks every button that is clicked and performs basic analyses on the text that students post (e.g., grade level or academic word use), inferential statistics were used to look for differences in students’ activities and writing habits. We primarily report on the types of activities where differences were observed. These activity types include the following:

- **Note**: This refers to a single post within the discussion forum.
- **Private shared note**: This is a post that is shared only with a specific peer or group of peers selected by the creator of the note. The peer or peers also have the ability to edit the note.
- **Messages to instructors**: These are private Pepper messages that are similar to e-mail.

- **Editing**: This refers to the number of revisions made to notes by those notes' creator. Changes could be as simple as correcting the conjugation of a verb or could include substantive changes to the content or subject of the posting.
- **Liking**: This refers to the number of Likes that a note has received. Likes indicate an acknowledgement that is similar to the Like feature on Facebook.
- **Links created by note**: This refers to the number of links to other notes compared to the number of notes created. Linking within Pepper is similar to tagging in social media.
- **Note rereading**: This refers to the number of times a note is viewed following the reader's first reading of that note.
- **Replies**: These are responses to other notes. In Pepper, replies are indented in a similar manner to those in a threaded discussion forum.
- **Sentiment**: This is a measure of the amount of emotion that is present in the note. Sentiment values are calculated using natural language-processing techniques (Fakhraie, 2011).

The analysis of student activity levels informed which types of student activity we modelled. An initial list of activity types was created by selecting all of the event types in which students' usage levels differed significantly by course facilitation method. Other activities were added when they had prior empirical or theoretical support, as was the case with liking (Phirangee & Hewitt, 2016). The list of activities was then reduced based on a correlational analysis: Items from the initial list that were strongly correlated ($r > .90$, $p < .01$) were identified, and a single item from that group was chosen. For example, notes and replies had a strong relationship ($r = .976$, $p < .001$), so we analyzed students' note-writing activities because this was their primary course activity, which also meant that there was more data from which the models could be built. The list was further reduced to those activities where sufficient data was present. For example, the linking behaviors of students were not modelled, because there were fewer than 50 events across all courses, which would have led to overfit models that would not generalize. This left four activity types: note writing, revisions, rereading, and Liking.

To prepare the data for modelling, the number of times that a student performed a specific activity within Pepper was calculated for each week within the year. These activity counts were then used as the model features (the attributes that the algorithm reasons over to create models). We then labelled each week's activity count to communicate when the activities took place and enable our interpretation of the resulting models. Weeks during the regular term were assigned a label that began with a "W" and ended with that week's number within the term. For example, activities from the third week of the term were assigned a label of "W3." Since activity levels were lower during the exam period and each course had different deadlines for its term-end paper, all of the student activities that were recorded during the exam period were combined. The exam period activities were then labelled with an "E." All of the activities that occurred after the exam period had ended were aggregated for similar reasons. These activities were then labelled with an "A" standing for "after." Like the activities that were logged during the exam period ("E") or after the term had ended ("A"), those that were logged more than a week before the beginning of term were aggregated. These activities took place in mid-to-late December and early January as instructors were preparing their courses and notifying students that the course materials had been posted or that the course was about to start. These events were labelled with a "P+" to indicate that they occurred well in advance of the beginning of the term. The activities that were logged in the week immediately prior to the start of term were labelled "P1." These counts were then used as inputs to the model and appear in the charts that illustrate the identified models.

Models of students' activities within Pepper were then identified using RapidMiner 5.3.015 (2014). These models represent students' system usage patterns. The *k*-means clustering algorithm was then applied to the normalized weekly usage statistics of individual students. The purpose of using *k*-means clustering was to create tight clusters or groups of students and to have each of these groups be as different as possible from each other. The *k*-means clustering was used because it has a history of being successfully employed to identify students' usage patterns within educational technology (Baker & Yacef, 2009; Brooks, Erickson, Greer, & Gutwin, 2014). Furthermore, it enabled us to explore which student groupings might be appropriate based on the similarities in their normalized weekly usage counts. Students' usage counts were normalized using *z*-transformation: This type of statistical normalization converts the data so that it fits a normal distribution that has a mean of 0. Students whose weekly activity counts match the mean are assigned a value of 0, those with activity counts above the mean are assigned an appropriate positive value, and those with activity counts below the mean are assigned a negative value. Performing this transformation helps ensure that extreme values do not unnecessarily influence the results of the modelling process. The results of clustering algorithms, including *k*-means, are also influenced by the data points that are selected as the cluster center points (Witten & Frank, 2005). To overcome this bias, precautions were taken: Each cluster's center point was randomly selected, and the process was repeated 1,000 times. Running the algorithm this many times with randomly selected center points reduces the likelihood that a poor initial choice will negatively affect the results.

While some clustering methods try to determine the appropriate number of clusters, *k*-means does not (Witten & Frank, 2005). As a result, we iteratively explored the number of clusters by changing the value of *k* (i.e., the number of clusters that the algorithm is expected to create). We started by setting *k* to 2, and increased *k* by 1 each time that we ran the analysis. We stopped adding more clusters once it became clear that we were creating groups whose activities were similar to one another. We used a combination of information to make this decision. We used the Davies-Bouldin Index (DBI), which measures the distance between individual clusters and the distance between each student who was assigned to a cluster. We also visually inspected the clusters and considered the number of students assigned to each cluster, the interview data, and educational theory to determine the appropriate value for *k*. We selected the value for *k* (i.e., the number of clusters) that minimized the DBI without creating unnecessarily small clusters. The *k*-values where a cluster contained only one or two students were kept only if there was strong theoretical support for their existence or the interview data confirmed the identified behavior pattern. The identified clusters were then assigned labels based on existing theory and the interview results.

The above analyses and modelling were considered in conjunction with interviews to further understand students' online experience and behavior. Semistructured interviews, an hour in length, were conducted with six students from the participating online courses. An e-mail inviting students to participate in an interview was sent to the six postgraduate courses, and those who expressed interest were selected to take part in the interview. Interviews were transcribed (please see Appendix A), and then thematic analysis was used to analyze each interview, which involved searching the data for patterns and themes to generate research insights about the phenomena (Glesne, 2011). After each student interview was analyzed, it was categorized by course facilitation method, which had been determined by the students themselves. This student-labeling method of course facilitation allowed us to compare both types of courses more accurately and generate themes that reflected students' online experience in each type of course.

Results

For this study, the quantitative data revealed only which activities students engaged in online, whereas the qualitative data provided students' understanding of their online behavior. Students from the instructor-facilitated (M = 33.62 hr, SD = 21.95) and peer-facilitated online courses (M = 27.16 hr, SD = 21.01) spent a comparable amount of time logged into the OLE ($p > .05$). Similarly, students' session counts from the peer-facilitated (M = 104.23, SD = 77.79) and instructor-facilitated online courses (M = 100.38, SD = 62.65) did not differ significantly ($p > .05$). The majority of the results focus on student activities; teacher activities are included only in the modelling of system activity patterns. These patterns, along with student interviews (i.e., their experiences, activities, and sense of community), are discussed below.

Students' Sense of Community and Usage

Table 1 (next page) shows the CCS scores for individual courses. It also details the facilitation method used within each course and the relationship between that facilitation method and the level (high or low) of community that was present within each online course. The tight relationship between the results of grouping the courses based on their CCS scores and the primary facilitation method that students identified as being used within their course led to our performing subsequent analyses from the perspective of the course's facilitation method. This small change in perspective allowed us to retain course C data for later analyses.

From Table 1, we can see that students' sense of community is higher ($F_{5,49}$ = 3.44, p = .010) in instructor-facilitated online courses (M = 55.75, SD = 10.60) than in peer-facilitated online courses (M = 45.67, SD = 9.07). This difference may be due to students' associating peer facilitation with a lack of instructor involvement. In particular, participants from both types of course expressed a preference for instructor facilitation but for different reasons. Students in peer -facilitated online courses (C, D, F) strongly emphasized that the instructor is needed to validate their ideas and thinking or to show an interest in student learning. As one participant stated,

But my preference is a bit of a balance, so the online courses that I've enjoyed are when the instructors will post a summary of the week, an introduction to think about some things, here's my voice adding to the conversation.

Another student also wanted the instructor more involved in the online course. She said,

I prefer hands-on. I think that allows the instructor to be with me in my learning journey. Also, when they're hands-on it seems like they're interested in what the class as a whole is learning and is studying, and I think that also develops the sense of community and deepens it because they are just more interested in what we're learning.

Conversely, students from instructor-facilitated online courses (A, B, E) felt that instructors are needed to keep the discussions on track by providing feedback. One participant stated,

> I really like what the professor did. He was there all the time, he was constantly showing his presence online, so I knew that he was there and he would reply to us…. He seemed like he was one of us, but he chimed in to make some relevant points, and he's clarifying, and he's not letting us go on and build on misunderstandings because there are errors, and he didn't want us to develop a fake understanding of something or a wrong one or any equities to be promoted or other things.

Similarly, another participant shared how much she enjoyed her professor being involved in the online discussion and disliked the "hands-off" professor. Reflecting on the facilitation methods this participant experienced in other online courses, she stated,

I like the hands-on, the facilitator. I just don't want the hands-off prof. I would consider my professor hands-on, like I said she was engaged, she gave really good feedback, she scheduled phone calls with us, which was really nice to give us feedback and talk to us about our final project, so that was a nice…. So, a prof who uses their experiences to be hands-on and stepping in when they need to guide, share or clarify. So yeah, being a facilitator when it comes to the discussion.

Table 1 CCS Scores by Course and Facilitator Type

Course	*n*	CCS *M*	CCS *SD*	Facilitator type	Community level
A	12	54.75	10.24	Instructor	High
B	8	57.38	3.24	Instructor	High
C	7	47.43	8.81	Peer	N/A[1]
D	14	47.36	8.50	Peer	Low
E	8	55.63	9.41	Instructor	High
F	6	39.67	9.56	Peer	Low

[1]The CCS score for course C did not differ significantly from any of the other courses.

In summary, all participants wanted the instructor to be involved in the online discussions. Instructor involvement reinforced the belief that the discussions would stay on track and that students would receive consistent feedback. More importantly, instructor participation allows students to feel as though the instructor values students' views and ideas and wants to be there.

Table 2 shows that there are significant differences based on the course's facilitation method for all student actions except Liking. It also shows that students from instructor-facilitated online courses posted more notes, performed more editing and linking actions, reread each others' notes more, and replied to a greater number of their classmates' notes.

In contrast, students in the peer-facilitated online courses were generally less active within the course, demonstrated slightly more sentiment in their prose, were more likely to communicate directly with their instructors, and participated in more private conversations.

Table 2 Descriptive and Inferential Statistics of Learner Use of the Pepper OLE

	Instructor-facilitated		Peer-facilitated		Mann-Whitney test	
Student activity	*M*	*SD*	*M*	*SD*	*U*	*p*
Notes	55.28	37.01	37.91	29.53	940.50	.009
Private shared notes	0.36	1.29	0.79	1.44	1,086.50	.025
Messages to instructor	1.17	2.57	3.77	4.84	949.00	.008
Editing	25.27	31.54	11.62	16.97	921.00	.007
Liking	0.89	0.73	0.57	0.33	1,241.50	.522
Links created by note	0.04	0.08	0.01	0.04	1,071.00	.023
Note rereading	198.21	151.51	123.56	121.50	889.00	.003
Replies	44.17	34.53	28.46	25.87	988.00	.022
Sentiment	6.37	0.14	6.48	0.15	736.00	<.001

Student Usage Patterns

We describe the models that were identified for each of the targeted activities: note writing, note editing, note rereading, and note Liking (please see Table 3). Each of these descriptions is accompanied by a visualization that represents the activity level of the typical student from the cluster that is associated with the identified usage pattern. The median activity levels of all students from the identified cluster were used to generate this representative student.

Table 3

The Number of Students and Instructors Assigned to Each Cluster for the Activities Where Usage Patterns Were Identified

	No. of people per cluster by course type		
Cluster label	*Instructor-facilitated*	*Student-facilitated*	*All*
Keeners	2	1	3
Outsiders	18	22	50
Respondents	17	22	39
Discussants	7	11	18
Perfectionists	5	1	6
Consistent Editors	4	3	7
Confident Writers	45	52	97
Keeners	1	1	2
Typical	21	46	67
Just-in-Time Readers	13	9	22
Outsiders	19	0	19
Maintainers	8	6	14
Nonsupporters	44	48	92
Supporters	2	2	4

Notes. The clustering algorithm identified four different usage patterns in students' posting of notes to the discussion forum (Figure 1):

- *Keeners* (n = 3): Members of this small group of instructors and students (see Table 3) had activity levels that were consistently higher than those of the other groups. The activity levels of these users often exceeded the combined activity of the other three groups.
- *Outsiders* (n = 50): These students performed the minimum activity required by their courses. Members of this group felt excluded from the community or opted out because they did not feel that being a part of a learning community was necessary or important. From Figure 1, we can see that they cease their participation during the final week of class (W13) and that they do not maintain their participation during the term break (W10).
- *Discussants* (n = 18): Members of this group initiate communication while the class is in session. They are the students who aim to start and maintain conversations
- *Respondents* (n = 39): Members of this group actively participate but may wait for others to start discussions. This is shown through their generally lower activity levels and the shape of their curve, which is similar to that of the Discussants but appears to be offset by about a week with activities continuing through the term break and into the exam period.

Interviewees from instructor-facilitated online courses (A, B, E) emphasized their awareness of their audience. Although they knew the professor would read and compare their notes to others, they focused more on contributing to the discussion by sharing their experiences and writing for their peers to connect with them and support their learning. For instance, one student stated, You want to let others know that you read their stuff, that you understood them and the topic, as such. But also you want to connect it with your experience to make it meaningful…. So what I try to do is connect the reading, the article with my experience as a person. Perhaps, as a husband, as a brother, also as a language teacher. So, you try to put together many things and hopefully one of my peers can connect with it.

Similarly, another participant echoed the importance of sharing personal experiences to help students connect and relate and, more importantly, learn from each other. She states,

What I like to do was make a personal connection to what we were reading. So I'm trying teaching because I want to make it more applicable and something other people can relate to.

The idea of connecting with peers was more important to students because it was believed that this connection would lead to deeper thinking and more meaningful posts. Interestingly, these students knew the professor would be reading their notes but opted to focus on building relationships to produce a deeper and collaborative online dialogue. As one participant stressed,

[When] I create my initial post, I know that the audience will be my professor and my peers…. I'm trying to look and add to what people have to say, or maybe there was something in their post that reminded me of something that somehow relates to the course material that I could actually go to and share.

Overall, students in the instructor-facilitated online courses focused on creating notes that were more meaningful by sharing a personal experience in order to connect with their peers. These students stressed not only their own learning but the importance of their peers' learning too.

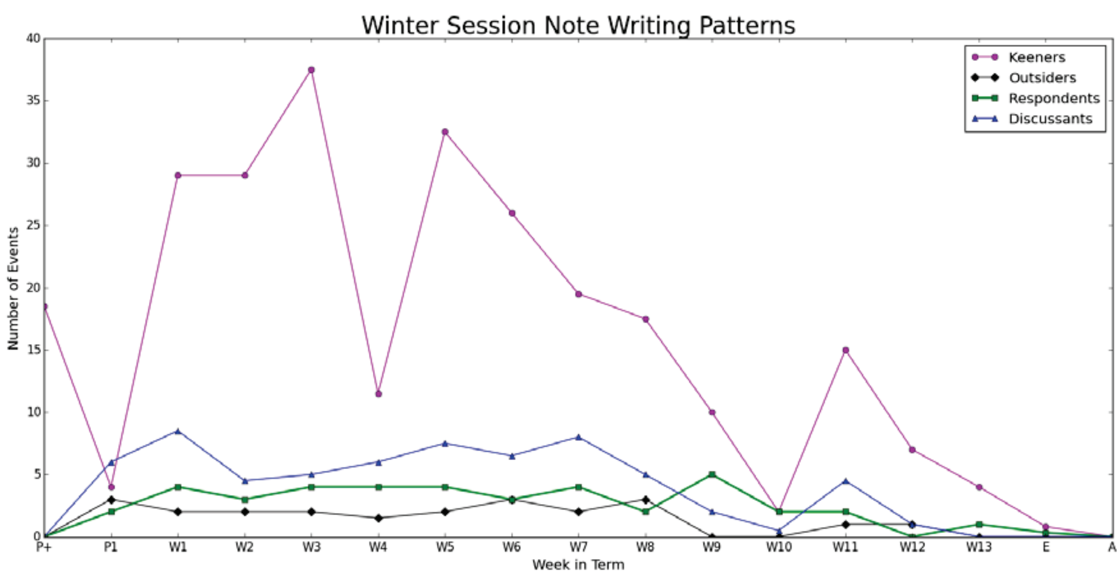


Figure 1. The number of new notes that each type of user typically wrote each week.

Although, students in the peer-facilitated online courses (C, D, F) expressed the importance of keeping the dialogue going and providing feedback to their peers, all admitted that they simply wanted to share their views and meet the requirements. More specifically, some of these requirements involved posting a certain number of notes and having to respond to all replies affiliated with those authored notes. One participant stated, [I'm] trying to express my views. But if I'm writing a note in response to someone I'm more focused on what they've said, their opinion, and if there's anything that I think would help like an article or something I'll attach it and be like, "Hey you'll enjoy this, it fits into what you're talking about." So, it really depends what kind of note it is…. But usually it's me putting my thoughts out there.

Similarly, another participant echoed the importance of helping her peers but focused on the need to express her own views and meet the requirements. She stated,

I try to fulfill what's required and express my perspective because I'm explaining what I believe about it, but if I'm replying [then] I'm giving feedback from my point of view, and also to help my peer like if they missed something but it's from my perspective…. I created my notes to meet the requirements the professor gave us. I didn't do more.

It seems as though students in the peer-facilitated online courses created notes in hopes of meeting the requirements, which included giving peer feedback and sharing their views on the course readings.

It is important to mention that some participants from other peer- and instructor-facilitated online courses witnessed some of their peers feeling left out or ignored within the online discussion. One participant from an instructor-facilitated online course recalls a time when he stopped commenting on a peer's post because he assumed she was not interested, but he discovered that she just felt her voice had no place in the discussion. He stated,

> She didn't participate as much as I expected…. She never answered me back and because of that I thought she was a distant person, so I stopped commenting on her post.... But then I get to meet her face-to-face and she's the opposite, she's a very kind and happy person, she's very sharing and the issue was that she felt overwhelmed.... She said, "Sometimes the professor would give some questions but people went of topic or they developed it in a certain way that I couldn't participate."

Similarly, a participant from the peer-facilitated online courses admits that she opted not to write more than the required amount of notes because she thought her peers would not understand or value her cultural experience. She stated,

> No, I wrote notes to meet requirements. It's not pleasant for me because I'm an ESL student, and considering my cultural background being from Kuwait, I don't know maybe I just didn't feel like people would get my experience. So, it's just like why would I want to spend so much time anyways. I could take these hours and do something else, instead of just reading what other people were saying [because] some of them might be interesting and some of them will be boring.

Therefore, some students in both peer- and instructor-facilitated online courses in this study are opting to limit their participation because they assume that their views will not be valued or accepted by peers, thus feeling like outsiders.

Editing. Only three note-editing patterns were identified (Figure 2): students who performed very few editing actions at the beginning of the term (the Confident Writers, $n = 97$), edited each post that they made once or twice (the Consistent Editors, $n = 7$), or performed extensive editing activities that in some cases exceeded their note posting activities (the Perfectionists, $n = 6$).

All interview participants, regardless of course type, revealed that they edited their notes, mainly to fix grammatical errors that they caught after posting. Participants from instructor-led online courses (A, B, E) seem to portray themselves as consistent editors. For example, one participant emphasized that it was extremely important to edit notes because of their permanency. She edited her notes “all the time because it’s permanent it’s different than an oral contribution. I’m doing my posts in the evening because I work, so I love the edit button. Thank God there’s an edit button.” Similarly, another participant found herself using the edit button regularly when writing posts because she knew others would be reading her work. She stated, “I post, and then I read it as if I’m the person who receives it, which is weird [laughs] and then I find something in it and I click the edit button and I re-edit.” One participant emphasized that he consistently edited his notes before and after posting because he would find mistakes. He stated,

> Yes, I do. It mainly because as English is my foreign language and I learn it as a foreign language sometimes I make spelling or typo mistakes. I’m a kind of perfectionist, so I like to check before and after I’ve posted it. If you check my notes, sometimes the due date was March 15, but then you see that I edited it in April, and it was because I reread it and found a spelling mistake like *our* instead of *or*.

Overall, participants from instructor-led online courses presented themselves as consistent editors, especially when the forums got busier throughout the term. Students attributed the need for this consistency to pleasing their audience (i.e., their instructor and peers) by presenting confident writing in their notes, whereas participants from peer-facilitated online courses minimally edited their notes and seemed to be more confident writers from the start.

Although students from peer-facilitated online courses (C, D, F) acknowledged editing their notes for grammatical errors and to adding missing content, many emphasized that they did so minimally and stopped once a peer replied to a post. For instance, a participant stated,

> Usually before I post anything I definitely check it. I do it first on a Word document and then post to Pepper. But after I post it I rarely go back to check it. So, usually I would revise any grammar or spelling mistakes.

Similarly, a participant argued that editing for grammar is allowed before or after but editing to add content, such as revising a paragraph, is wrong. She stated,

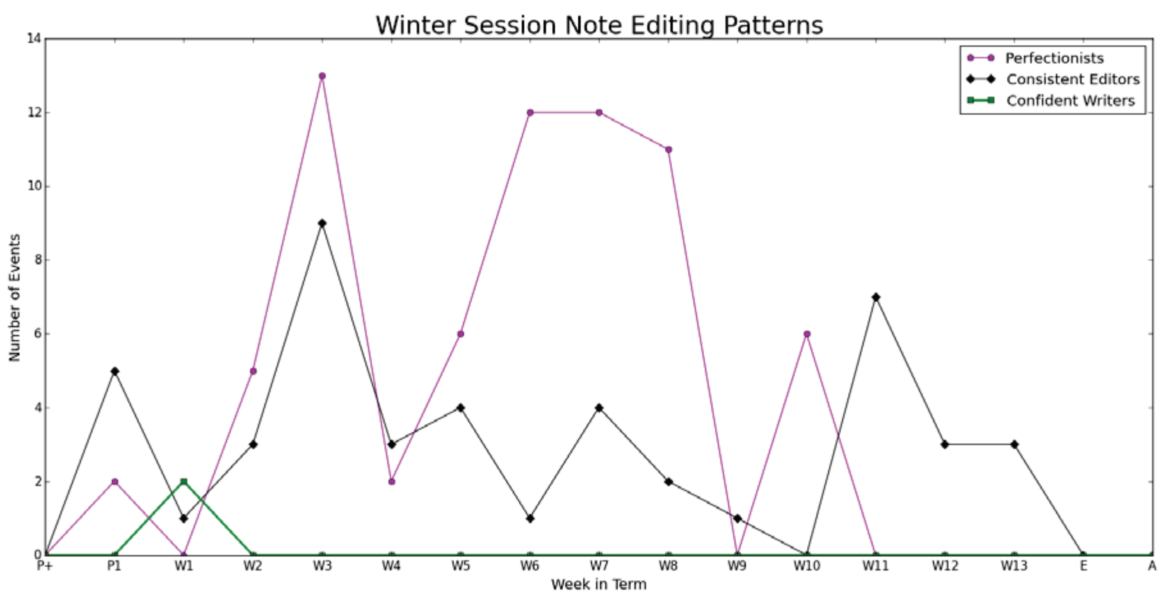


Figure 2. The number of weekly edits typically made by each user type to existing notes.

I don't really revise them because I feel like that's cheating. If you've written a note [pause]. If it's a typo you're fixing that's one thing, but if you've written a note [trails off]. I find this frustrating like someone wrote a note and I responded to it, and then they revised it and included a paragraph, which I commented on, then replied to me saying, "Oh, well what did you think about this?" And so, I just looked like an idiot because my comment made no sense anymore, so I think there are problems in there sometimes. So, yeah there was a whole other section they put in when I said, "Hey, have you thought about this side of the problem? I'm sure that would be a really conversation." So, I try to stay away from revising posts after that because it was just a strange experience…. I rather just write what I had to say knowing that I got gaps.

Overall, participants from peer-facilitated online courses edited their notes minimally for grammar and content compared to those from instructor-facilitated online courses. This behavior may be due to students knowing that their instructor was not observing their discussions consistently and their feeling more comfortable posting their notes with gaps. This finding is supported by previous research in which students in peer-facilitated courses were more likely to vocalize their views, challenge each other, and negotiate ideas (Baran & Correia, 2009; Rourke & Anderson, 2002). Lastly, it is important to recognize that despite course type, students stopped editing their note when a peer replied. The reason for doing so came from a desire to not disrupt the flow of the online conversation and to respect their peers' comments.

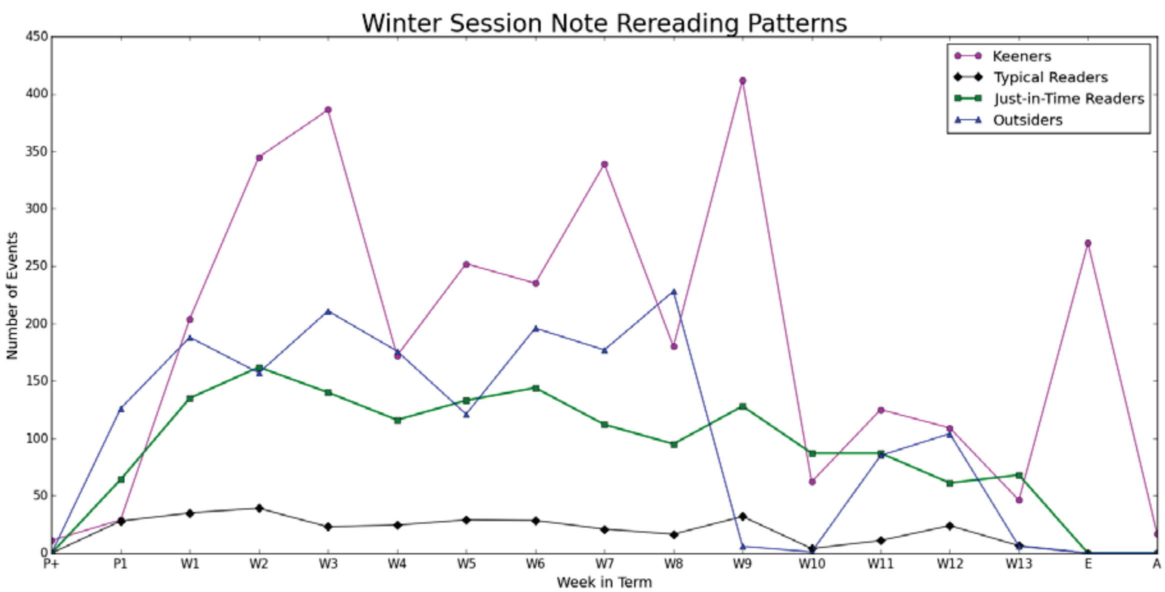


Figure 3. The typical number of note rereads that each user type performed.

Rereading. The rereading of a note by members of the community suggests its influence on the community or its importance to members of the community. As with note posting, two of the four rereading patterns that were identified (Figure 3) were those of the Keener (n = 2) and the Outsider (n = 19). In this case, the activity of Keeners extended beyond the term, which may represent attempts to understand instructor feedback in light of the course discussion or instructor grading activities. The Outsider rereading pattern was also different: Their reading activity levels were far higher than their posting activities. Outsiders reread instructor posts multiple times to try to understand instructor-provided summaries and the instructor's expectations. Outsiders also reread students' posts, possibly in an attempt to better understand postings where cultural references had hindered understanding, which may be why the shape of their curve resembles that of the Keeners and why the cessation of their rereading activities coincided with the last week of class. The third group of students are the Typical Readers (*n* = 67), who did relatively little rereading. The final group was the Just-in-Time Readers (*n* = 22), whose consistent level of rereading seemed to follow course demands, continued through the term break, and ended with the exam period.

For participants in the instructor-facilitated online courses (A, B, E) rereading was seen as a key behavior in understanding and connecting with a peer's note. For these participants, rereading was done as an audience member to make sure their notes accurately and clearly reflected their thoughts. For instance, one participant stated she would write her note then leave it and return to it another day with fresh eyes. She shared her method:

> I would also go back in the morning, so I would do a Word document and then go back in the morning "Does this make sense?" Rereading for errors because it's something that's permanently going to be in cyberspace for everybody, and you can't take it back. So, yes I do reread often.

In the same way, another participant stated,

> To be honest, I did it a few times. It was mainly because I was preparing an answer, a reply for them, and I wanted to make sure to cover all the points and to not misinterpret what said but apart from this I didn't do it.

The idea of rereading as an audience member highlights a strategy these students used to avoid misunderstandings, provide accurate feedback, and show their peers that their work was valued. Perhaps adopting this particular rereading method produced more confident writers because students were able to reflect and synthesize their ideas before publicly posting them to the online discussion forum, which confirms Hewitt's claims (2005) that "the asynchronous nature of the interaction allows learners to reflect in greater depth before they share their ideas publicly" (p. 568). This is key to building confident writers, because it provides students with an opportunity to reread, reflect, and rewrite if needed. Another participant emphasized that it is important to reread because this leads to proper communication and a stronger connection in the forum, which is necessary since online platforms are missing various social cues. For her,

> So, I'm rereading as an audience member and one of my peers, and if you think about it, it's written communication and because you're missing other cues and the other face-to-face stuff, I found myself making sure it was polite enough by how it sounds so that it's not offensive. So, it's very important for me to pay attention to the way somebody who receives it might read it.

Students from the instructor-facilitated online courses reinforced the notion of rereading as an audience member, as a final proofing stage, to make sure they replied accurately without any misunderstanding. These students understood that such communication was critical in shaping and building a dialogue with their peers and made the effort to reread their notes before posting them publicly. Perhaps knowing and seeing the instructor's involvement in the course also motivated students to be aware of what they write and how they respond, hence their time and effort in rereading notes.

In contrast, students in peer-facilitated online courses (C, D, F) did little rereading. Their activities consisted of rereading their own notes to make sure there were no errors and rereading the instructor's notes to make sure they were meeting the requirements. As one participant stated,

> I reread because when I write it initially [pause] sometimes the sentence makes sense but it's out of context, so I just fix it by rephrasing. So, I do it to make sure it was clear, that I didn't make any mistakes here and there.

Another participant stated, “No, I didn’t read my peers ever [laughs]. I might have gone back to reread if something that I had written before might have been relevant or something that I had read before was relevant.” There is a focus from these students to reread only their notes rather than their peers’. If they did reread a peer’s note, it was because it contributed to their own learning. For instance, another participant built on the idea of rereading because it was the professor’s notes and it was beneficial to her work. She states,

> I’ve reread notes. I do it all the time, especially the professor’s notes, but then sometimes I’ll just remember a really good conversation. And because I’m working as well, if I think there’s something in there that could really relate to what I’m doing at work or I think I should remember, then I’ll go back and look at it.

Perhaps knowing that the instructor is not fully involved in the online course indirectly created such behaviors as minimally rereading and only doing so when it benefited students’ own learning or if they had to. In comparison to students from the instructor-led online courses whose main purpose for rereading from the viewpoint of their audience was to support and benefit their peers, students from the peer-facilitated online courses reread less and did so only to benefit themselves.

Liking. Similar to editing, three patterns of note Liking were identified (Figure 4). The Nonsupporters (*n* = 92) are similar to the outsiders that were found when analyzing students’ and instructors’ rereading and note-posting habits: They were minimally active and did not attempt to support their classmates through the use of Liking. The Maintainers (*n* = 14) are typical students who are reciprocating the activities of others and trying to meet their social obligations, whereas the Supporters (*n* = 4) aim to give others the sense that they have contributed to the learning community. Supporters try to impart this sense of contribution by liking others’ notes, which serves to acknowledge the contribution that is made by the note’s author.

Participants admitted to using the Like button minimally when engaging in online discussions. Those in the instructor-facilitated online courses (A, B, E) were eager to emphasize the importance of both the instructor and students using the Like button. As one student indicated, Liking validated the work she was sharing and that her peers valued her ideas.

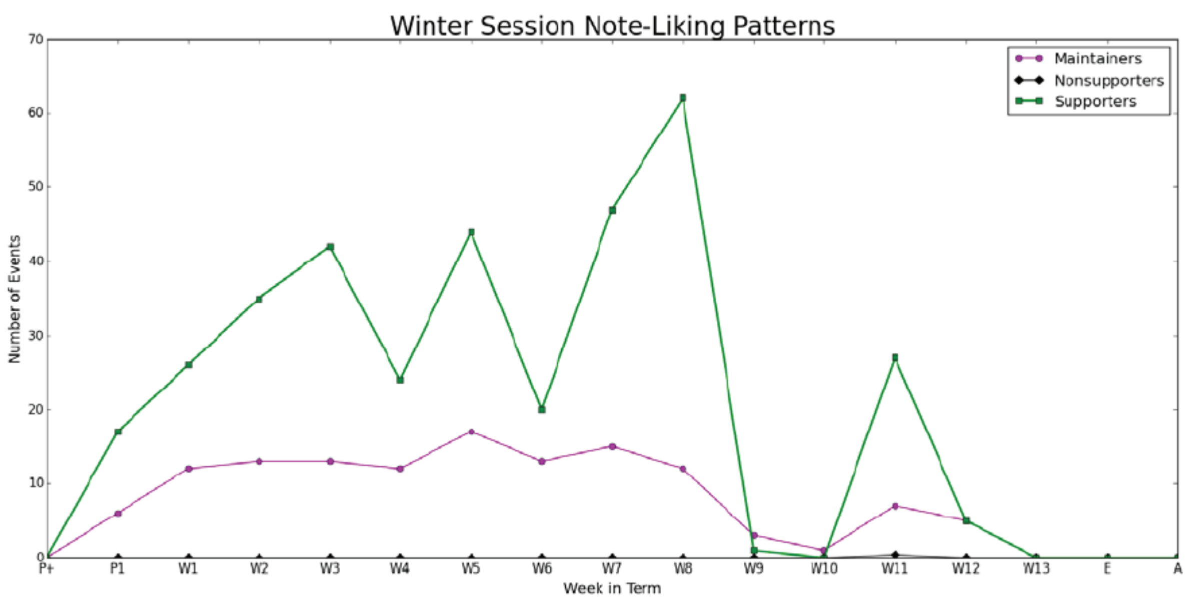


Figure 4. The number of likes that each user type gave.

She states,

> Also, do you know how important the Like feature is? I look at who Liked my note. And I'm thinking, "Oh wow somebody Liked my note." And so, I found myself looking for how many Likes I got, or if somebody Liked my note because it validated that I was on track, or that they Liked my stuff.

Another participant highlighted that when the instructor used the Like button, it reinforced their support and validation for students' ideas and views. As another participant expressed,

> I really like when the instructor answers your notes. Even to say "Like," you know using the Like button, or just to chime in and say good job because you know that she cares about you. I mean the instructor is actually reading your work.

Perhaps having an instructor who oversees the online discussions and peers who write for each other diminishes the use and purpose of the Like button in this type of course. Yet, for many of these students, the Like button is a useful feature that supports and validates their work in a quick and simple way.

Participants from peer-facilitated online courses (C, D, F) also used Liking to simply acknowledge a peer's contribution rather than write a response. As one participant summarized,

> Sometimes it was just a really good conversation that other people were having, so I would be Liking some of those because they were really good, and I was thinking, "You guys have this covered." Maybe you don't log on for a few days and you see that. But they just had an amazing conversation and you learnt a lot from that, and I don't necessarily have anything to add, so I'd just Like it.

These students seem to use the Like button as a social cue to let their peers know that they had read a note, because they did not submit a formal reply. Not having an instructor consistently involved in the online discussions seemed to motivate students to strategically use the Like button to acknowledge their peers' work, thus meeting their social obligation of acknowledging the conversations. Overall, students in both course types used the Like button to support their peers' work. However, it held more meaning for students in the instructor-facilitated online courses; they emphasized that this button is needed to provide quick and continual validation of their work from their peers and the instructor, which confirms previous research (Phirangee & Hewitt, 2016. Therefore, Liking seems to play an essential role in fostering a sense of community within OLEs.

Discussion and Conclusion

Recent research suggests that students prefer instructor-facilitated online courses (Hew, 2015), despite the reported benefits of peer facilitation. Like Hew's (2015) students, those from this investigation preferred instructor facilitation because instructors are the "subject matter experts," are able to keep discussions on topic and ensure equity, and are better at guiding learning. Lastly, students prefer them because they want instructors to facilitate discussions rather than act as sages; students expressed a need, want, and preference for instructor facilitation regardless of the facilitation method that had been used in their course. Moreover, the developed models detail how this preference is manifested through student behaviors. This finding supports Shea, Li, and Pickett's (2006) research, which found a strong connection between teaching presence, learning, and community:

> The respondents to the survey were significantly more likely to report higher levels of learning and community when they also reported that their instructors exhibited more salient "teaching presence" behaviors…. This study reveals that a strong and active presence on the part of the instructor—one in which she or he actively guides and orchestrates the discourse—is related both to students' sense of connectedness and learning. (pp. 184–185)

Additionally, some studies found that instructors needed to stay involved to promote collaboration and conversations among students in OLEs because higher levels of engagement and interactions among students lead to deeper and more critical thinking (Zach & Agosto, 2009; Agosto, Copeland, & Zach, 2013). Furthermore, in the community of inquiry (CoI) model, teaching presence is presented as one of the needed elements for building a community in an OLE. This element emphasizes that the teacher is needed to do the following: design and establish the learning experience before and during the course; teach, which involves implementing activities to encourage interaction between students, groups, the content, and teacher; and move beyond the moderator role and provide subject matter expertise through direct instruction (Anderson, Rourke, Garrison, & Archer, 2001). Therefore, instructors fulfill an important role in developing a sense of community in online courses. Perhaps deep background knowledge in a discipline is simply a necessary prerequisite for effective facilitation. Instructors can add content and perspective that adds value to discussions and makes them more engaging. Students tend to lack background knowledge in the subject matter, and their peers probably do not trust what little knowledge they do have. This is probably why peer-facilitated online courses were fundamentally less effective.

The quantitative data suggest that higher community scores, as measured by Rovai's (2002b) classroom community scale, were associated with instructor-facilitated online courses, whereas peer-facilitated online courses had lower community scores. Although students from both types of courses exhibited similar amounts of system use (i.e., hours spent online and number of sessions), specific student activities revealed that online behavior differed significantly between peer- and instructor-facilitated courses. Overall, data analytics showed that students in peer-facilitated online courses were privately communicating with each other more often, possibly relating to behind-the-scenes prep work for assignments or discussion moderation, whereas students in the instructor-facilitated online courses had higher participation levels in the online discussions, with more notes, replies, edits, rereads, and links created to other notes. Perhaps the higher levels of student activity in the instructor-facilitated online courses is partially due to a sense of constant surveillance. Students in instructor-facilitated online courses may expend more effort because there is a greater sense that the instructor is constantly present, which affects learner behavior. Students are more likely to engage in class discussions because they know the instructor is personally involved in those interactions and monitoring what everyone says. Interestingly, the cluster analysis also revealed that for notes, edits, rereads, and Liking, students in instructor-facilitated online courses were always thinking about their peers' learning, in hopes of connecting with them. However, those in peer-facilitated online courses focused more on themselves and their own learning.

Based on the developed analytics, an instructor-facilitated online course is more likely to stimulate student participation and build a stronger sense of community. These findings advance Hew's (2015) research by "measur[ing] the impact of peer or instructor facilitation on student outcomes such as the number of their postings" (p. 36). Our data provided a clear suggestion as to which online facilitation strategy instructors should adopt, when designing a course, to foster a sense of community in an OLE. Should an instructor want to employ peer facilitation, the developed analytics could be used to identify students who may feel excluded. The instructor could then intervene to bring those students into the community. This approach and use of analytics may help instructors and students to mitigate or overcome some of the risks that appear to be associated with peer facilitation. The use of these analytics by instructors as an early warning system and the study of its relationship to student outcomes is a reasonable next step, as is tracking the relationship between student activity patterns and the development of students' sense of community over the course of a term. Lastly, although instructors and graduate students use Pepper widely at our academic institution, it is not used across a variety of disciplines or institutions, which may limit the generalizability of our findings.

References

Allen, I. E., & Seaman, J. (2013). *Changing course: Ten years of tracking online education in the United States*. Newburyport, MA: Sloan Consortium.

Agosto, D. E., Copeland, A. J., & Zach, L. (2013). Testing the benefits of blended education: Using social technology to foster collaboration and knowledge sharing in face-to-face LIS courses. *Journal of Education for Library and Information Science, 54*(2), 94–107.

Angelino, L. M., Williams, F. K., & Natvig, D. (2007). Strategies to engage online students and reduce attrition rates. *Journal of Educators Online*, *4*(2), n2.

Arend, B. (2009). Encouraging critical thinking in online threaded discussions. *The Journal of Educators Online, 6*(1), 1–23.

Hewitt, J. (2005). Toward an understanding of how threads die in asynchronous computer conferences. *The Journal of the Learning Sciences, 14*(4), 567–589.

Baker, R. S. J. D., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining, 1*(1), 3–17.

Baran, E., & Correia, A. P. (2009). Student-led facilitation strategies in online discussions. *Distance Education*, *30*(3), 339–361.

Belcher, A., Hall, B., Pressey, K., & Kelley, K. (2014). An analysis of faculty promotion of critical thinking and peer interaction within threaded discussions. *Online Learning, 19*(4), 37–44.

Bolliger, D. U., & Inan, F. A. (2012). Development and validation of the online student connectedness survey (OSCS). *The International Review of Research in Open and Distance Learning*, *13*(3), 41–65.

Brooks, C., Demmans Epp, C., Logan, G., & Greer, J. (2011). The who, what, when, and why of lecture capture. In *International Conference on Learning Analytics and Knowledge (LAK)* (pp. 86–92). Banff, Canada: ACM.

Brooks, C., Erickson, G., Greer, J., & Gutwin, C. (2014). Modelling and quantifying the behaviors of students in lecture capture environments. *Computers & Education*, *75*, 282–292.

Correia, A. P., & Davis, N. E. (2007). The design of collaboration in the virtual classroom. *30th annual proceedings of selected papers on the practice of educational communications and technology, 2*, 84–87.

Fakhraie, N. (2011). What's in a note? Sentiment analysis in online educational forums (Unpublished master's thesis). University of Toronto. Retrieved from https://tspace.library.utoronto.ca/bitstream/1807/30086/1/Fakhraie_Seyedeh_Najmeh_201111_MA_thesis.pdf

Ferguson, R., & Buckingham Shum, S. (2012). Towards a social learning space for open educational resources. In A. Okada, T. Connolly, & P. J. Scott (Eds.), *Collaborative learning 2.0: Open educational resources* (pp. 309–327). IGI Global.

Bull, S., Greer, J. E., McCalla, G. I., & Kettel, L. (2001). Help-seeking in an asynchronous help forum. In *International Conference in Artificial Intelligence in Education (AIED)*. Retrieved from http://users.sussex.ac.uk/~bend/aied2001/bull.pdf

Carr, S. (2000). As distance education comes of age, the challenge is keeping the students. *The Chronicle of Higher Education, 4*(23), A39–A41.

Clarke, L. W., & Bartholomew, A. (2014). Digging beneath the surface: Analyzing the complexity of instructors' participation in asynchronous discussion. *Online Learning*, *18*(3).

Conrad, D. (2005). Building and maintaining community in cohort-based online learning. *Journal of Distance Education*, *20*(1), 1–20.

Correia, A. P., & Baran, E. (2010). Lessons learned on facilitating asynchronous discussions for online learning. *Educação, Formação & Tecnologias-ISSN 1646-933X*, *3*(1), 59–67.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage Publications.

Dzubinski, L. (2014). Teaching presence: Co-creating a multi-national online learning community in an asynchronous classroom. *Online Learning*, *18*(2), 1–16.

Ertmer, P. A., Richardson, J. C., Belland, B., Camin, D., Connolly, P., Coulthard, G., Lei, K., & Mong, C. (2007). Using peer feedback to enhance the quality of student online postings: An exploratory study. *Journal of Computer-Mediated Communication, 12*(2), 412–433.

Fowler, F. J. (2009). Survey research methods (4th ed.). Thousand Oaks: Sage Publications.

Gallagher-Lepak, S., Reilly, J., & Killion, C. (2009). Nursing student perceptions of community in online learning. *Contemporary Nurse: A Journal for the Australian Nursing Profession, 32,* 133–146.

Glesne, C. (2011). *Becoming qualitative researchers: An introduction.* Boston: Pearson Education, Inc.

Gray, D. (2013. Barriers to online postsecondary education crumble: Enrollment in traditional face-to-face courses declines as enrollment in online courses increases. *Contemporary Issues in Education Research (CIER), 6*(3), 345–348.

Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research*, *17*(4), 397–431.

Hew, K. F. (2015). Student perceptions of peer versus instructor facilitation of asynchronous online discussions: Further findings from three cases. *Instructional Science*, *43*(1), 19–38.

Hiltz, S. R., & Arbaugh, J. B. (2003). Improving quantitative research methods in studies of asynchronous learning networks (ALN). In J. R. Bourne & J. C. Moore (Eds.), *Elements of quality online education. Practice and directions* (pp. 59–72). Needham, MA: Sloan-C.

Hulon, S. (2013, March). Face-to-face discussions versus online threaded discussions: Can we have the best of both worlds? In *Society for Information Technology & Teacher Education International Conference* (Vol. 2013, No. 1, pp. 577–582).

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. In *International Conference on Learning Analytics and Knowledge (LAK)* (pp. 170–179). New York: ACM.

Larson, B. E., & Keiper, T. A. (2002). Classroom discussion and threaded electronic discussion: Learning in two arenas. *Contemporary Issues in Technology and Teacher Education*, *2*(1), 45–62.

Light, P., Nesbitt, E., Light, V., & White, S. (2000). Variety is the spice of life: Student use of CMC in the context of campus-based study. *Computers & Education, 34*(3), 257–267.

Liu, X., Magjuka, R. J., Bonk, C. J., & Lee, S. (2007). Does sense of community matter? *Quarterly Review of Distance Education*, *8*, 9–24.

Mazzolini, M., & Maddison, S. (2007). When to jump in: The role of the instructor in online discussion forums. *Computers & Education, 49*(2), 193–213.

Nully, D. D. (2008). The adequacy of response rates to online and paper surveys: What can be done? *Assessment & Evaluation in Higher Education, 33*(3), 301–314.

Ouzts, K. (2006). Sense of community in online courses. *Quarterly Review of Distance Education, 7*, 285–296.

Phirangee, K., & Hewitt, J. (2016). Loving this dialogue!!!! ☺: Expressing emotion through the strategic manipulation of limited non-verbal cues in online learning environments. In S. Y. Tettegah & M. P. McCreery (Eds.), *Emotions, technology, and learning* (pp. 69–85). New York: Elsevier.

RapidMiner. (2014). (Version 5.3.015). RapidMiner Inc. Retrieved from https://rapidminer.com

Richardson, V. (2003). Constructive pedagogy. *Teachers College Record, 105*(9), 1623–1640.

Rourke, L., & Anderson, T. (2002). Using peer teams to lead online discussions. *Journal of Interactive Media in Education, 1*, 1–21

Rovai, A. P. (2002a). Building sense of community at a distance. *The International Review of Research in Open and Distance Learning, 3*(1).

Rovai, A. P. (2002b). Development of an instrument to measure classroom community. *Internet and Higher Education, 5*, 197–211.

Rovai, A. P., & Wighting, M. J. (2005). Feelings of alienation and community among higher education students in a virtual classroom. *The Internet and Higher Education, 8*(2), 97–110.

Shackelford, J. L., & Maxwell, M. (2012). Sense of community in graduate online education: Contribution of learner to learner interaction. *International Review of Research in Open & Distance Learning, 13*(4).

Shea, P., Li, C. S., & Pickett, A. (2006). A study of teaching presence and student sense of learning community in fully online and web-enhanced college courses. *The Internet and Higher Education, 9*(3), 175–190.

Song, L., & McNary, S. W. (2011). Understanding students' online interaction: Analysis of discussion board postings. *Journal of Interactive Online Learning, 19*(1), 1–14.

Swan, K. (Ed.). (2009). Threaded discussion. *Encyclopedia of Distance Learning* (2nd ed., Vols. 1–4). Kent State University: IGI Global.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Witten, I., & Frank, E. (2005). *Data mining: Practical machine learning tools and techniques* (2nd ed.). Boston, MA: Morgan Kaufman.

Zembylas, M. (2008). Adult learners' emotions in online learning. *Distance Education, 29*(1), 71–87.

Appendix A - Transcription Legend

Mark-up	Purpose
…	Content was deleted.
[ ]	Content inserted to clarify or convey participants' body language.
Italics	Highlights the content participants' emphasized during the interview.
""	To illustrate that participants are quoting a peer or discussing their own thoughts while they were in the situation at the time.

Developing Learning Analytics Design Knowledge in the “Middle Space”: The Student Tuning Model and Align Design Framework for Learning Analytics Use

Alyssa Friend Wise, Jovita Maria Vytasek, Simone Hausknecht and Yuting Zhao
Simon Fraser University

Abstract: This paper addresses a relatively unexplored area in the field of learning analytics: how analytics are taken up and used as part of teaching and learning processes. Initial steps are taken towards developing design knowledge for this “middle space,” with a focus on students as analytics users. First, a core set of challenges for analytics use identified in the literature are compiled. Then, a process model is presented for conceptualizing students’ learning analytics use as part of a self-regulatory cycle of grounding, goal-setting, action and reflection—the *Student Tuning Model*. Finally, the *Align Design Framework* is presented with initial validation as a tool for pedagogical design that addresses the identified challenges and supports students’ use of analytics as part of the tuning process. Together, the framework’s four interconnected principles of *Integration, Agency, Reference Frame* and *Dialogue / Audience* provide a useful starting point for further inquiry into well-designed learning analytics implementations.

Introduction

Information derives its importance from the possibilities of action
(Postman, 1985, p68)

This paper addresses a relatively unexplored area in the field of learning analytics: how analytics are taken up and used as part of teaching and learning processes. Initial work in learning analytics has focused on how to capture, process, and present large quantities of data to educational stakeholders in useful ways (Verbert, Duval, Klerkx, Govaerts, & Santos, 2013; Romero & Ventura, 2013; Baker & Inventado, 2014). The result has been the development of a set of increasingly sophisticated tools for monitoring, predicting and, in some cases, assessing student activity (M'hammed, Abdous, & Yen, 2012; Greller & Drachsler, 2012; Verbert, et al. 2014; Vatrapu, Teplovs, & Fujita, 2011) as well as systems that make recommendations based on the data collected (Manouselis, Drachsler, Vuorikari, Hummel, &

Koper, 2011). However, the creation of these analytics is only half of the endeavor; in order for the data to actually influence learning process, these analytic *outputs* need to become *inputs* into subsequent decision-making (Clow, 2013). This latter activity of using learning analytics to inform choices and subsequent action in-situ is the focus of this article.

The study of how learning analytics are taken up and used in practice is important because this activity is decidedly non-trivial. History shows that the use of educational innovations (and designed objects more generally) is never fully determined by the form of the technologies themselves. Rather it is dynamically shaped by the affordances of the new tools in combination with the needs and abilities of the users and the constraints of the situations into which they are placed (Cuban, 1986, 2001; Gibson, 1977; Norman, 2013). Thus it is important to develop robust conceptualizations of the processes by which different users (policy makers, administrators, designers, instructor, students) can work with learning analytics in their particular contexts of action. In particular, consideration of the interplay of analytics in the relationship between teachers, students and learning environment addresses a very practical need: if instructors don't see how learning analytics can become a productive part of their classroom practices, they will use them only in an ancillary role, if at all. Conversely, through analytics, students, teachers and others have the potential to receive new forms of feedback as part of their educational practice. This feedback can be used to engage in new activities related to tracing progress and reflecting on the learning process, in both summative and formative ways, individually and as a group (Visser, Plomp, Amirault, & Kuiper, 2002; Macfadyen & Dawson, 2010). Without thoughtful design to encourage and shape analytics usage, these potential benefits will be lost.

This paper takes initial steps towards developing design knowledge for this "middle space," between analytics data presentation and action based on the analytics, with a focus on the case of students as analytics users. We begin by compiling a core set of challenges for analytics use from the literature. We then present a process model for conceptualizing students' learning analytics use as part of a self-regulatory cycle: The *Student Tuning Model*. Finally, we propose and conduct initial validation of a set of principles for pedagogical design that addresses the identified challenges and supports students' use of analytics as part of the tuning process: The *Align Design Framework.*

Literature Review

Context in the Field

As the field of learning analytics matures, the experiences and needs of educational actors (teachers, students, designers, administrators) who use analytics in their practice are emerging as important areas for research. The set of existing frameworks for learning analytics takes a decidedly researcher / developer viewpoint. For example, the commonly cited Campbell and Oblinger (2007) five-step model of learning analytics (Capture, Report, Predict, Act, Refine) is heavily focused on how to work with data, with the action in step four happening as an external intervention into the regular operations of the teaching and learning system. Chatti and colleagues (2012) take a step towards considering the eventual context of analytics use by introducing four sensitizing questions (What kind of data is used? Who is the target of the analysis? Why is the data being collected? How does the system perform the analysis?) but these questions are still meant to inform the design of the analytics system by the developers.

Clow (2013) proposes some initial steps towards thinking about a model of learning analytics use by highlighting the importance of "closing the loop" and considering the potential impact on the learners who initially generated the data. Specifically, drawing on the theories of Schön (1983) and Kolb (1984), he suggests that learning analytics use can be productively thought of as part of a reflective practice cycle in which the information provides feedback on the teaching and learning activities that can be used to adjust or experiment with changes in these activities. He also goes a step further to expand the notion of reflective practice as an individual activity to one in which analytics can facilitate conversation between teachers and students. A similar notion of analytics as part of reflective practice has been proposed by Brooks et al. (2014) in their "data-assisted approach to building technology-enhanced learning environments" (p. 123). This work is useful in providing a general theoretical framing for learning analytics use; however, a specific vision of what the learning-analytics reflective practice cycle might look like and how it can be supported remains to be provided.

Empirically, there has been some recent research into the potential for analytics to be used by teachers as part of their reflective practice (Ghislandi & Raffaghelli, 2015; Avramides, Hunter, Oliver, & Luckin, 2014; Melero, Hernández-Leo, Sun, Santos, & Blat, 2015; McKenney & Mor, 2015), as well as the development of specifically designed analytics tools that are embedded to assist teacher inquiry into student learning (Haya, Daems, Malzahn, Castellanos, & Hoppe, 2015). For example, the information provided though analytics can be used as data to support teachers in making decisions about when to intervene and support their students (van Leeuwen, 2015). These studies all underscore the idea that the use of a combined approach of learning design, teacher inquiry into student learning and learning analytics can produce effective new pedagogies. However, as Rodríguez-Triana, et al. (2014) point out, to take advantage of this potential, educators need to craft their educational activities in ways that maximize the use of the learning analytics. Lockyer, Heathcote & Dawson (2013) have proposed one way for doing so by developing a conceptual model for teacher's use of learning analytics as part of learning design.

In contrast to this attention to learning analytics use by teachers, there has been limited research exploring how students interpret and act on learning analytics. A small number of studies have started to examine the different kinds of reactions students have to particular types of visualizations (Santos, Govaerts, Verbert, & Duval, 2012; Arnold & Pistilli, 2012; Corrin & de Barba, 2014; Aguilar, 2015), but have yet to examine how students use these analytics in practice. A separate body of CSCL literature has worked extensively with student-facing analytics in the form of "group awareness" tools (e.g. Buder, 2011; Janssen & Bodemer, 2013); however, again the focus has been more on what group process information is provided and how (i.e. the design of the tools) rather than understanding and supporting students' processes surrounding their use. In sum, there is scant guidance in the literature about how to encourage and guide students in using analytics to support their learning. A more detailed consideration of the learner perspective is needed because this is an important audience for analytics use and one that has unique and different needs from teachers (Ferguson, 2012). We begin our attention to this issue by first examining the literature for core challenges in the use of learning analytics considered particularly from the perspective of learners.

Challenges in Learning Analytics Usage

At its core, the process of using analytics to inform teaching and learning is comprised of two central activities: making sense of the information presented in the analytics and taking action based on this information (Siemens, 2013; Clow, 2013). Sense-making involves self-evaluation, asking useful questions of the data and finding relevant answers (Buckingham Shum, 2012; Verbert et al, 2013). Taking action relates to what we do (differently) based on this information; cognitively this is about decision-making. Overall, the act of interpretation has received greater attention than that of decision-making; however, we expect that as attention to the process of analytics use increases, this will become more balanced.

Below we review some of the central challenges that learners face in working with analytics, compiled from issues previously identified in the literature, addressing both the act of interpretation and subsequent action.

Learner interpretation.

Interpretation: The Challenge of Context. Interpretation refers to the link between the presentation of analytics to learners and how the learner makes meaning from this information. One of the fundamental challenges when it comes to interpreting analytics is providing context, as analytics are inherently detached representations of past activity. Giving them meaning requires a consideration of where the data is drawn from, why it is important and what role it should plays in future decisions. This is a complex process. From a teaching perspective, "knowledge of actual course design and instructor intentions is critical in determining which variables can meaningfully represent student effort or activity, and which should be excluded" (Macfadyen & Dawson, 2010, p. 597). Similarly, for learners, students need to be able to contextualize what is represented by these analytics, as well as what that means in relation to their personal goals and the overall intentions of the course. In many cases it seems to be assumed that simply providing well-designed analytics will be enough to induce productive use. However, there are several factors that work against this. One particular concern is that students are often not privy to their instructor's pedagogical intentions, and thus unaware of both the learning goals for an educational activity and what productive patterns of engagement in it (as indicated by the analytics) would look like. Additionally, without guidance, students may not have the metacognitive skills needed to understand how analytics can support self-monitoring and/or reflective activities (Butler & Winne, 1995). Thus, students may be unclear as to what questions to ask of the data and how it can be used to support their learning.

Interpretation: The Challenge of Trust. A second challenge in the interpretation of analytics from the learner perspective is trust in the fidelity and usefulness of the analytics presented. If students don't believe that the analytics are collecting and presenting data that are accurate and useful for their needs, they won't spend time interpreting them. Social acceptability of the analytics depends on trust (Clow, 2013; Macfadyen & Dawson, 2012) and without that, analytics can have no influence on the learning activity.

Additionally, to have trust there needs to be a level of understanding. With the ubiquity of data tracking, most students are aware of the fact that actions in digital spaces leave trace data. The harder conceptual leap is that these remnants of complex actions can be extracted and analyzed to reconstruct higher level representation of their actions. Students are often unaware of how they are being monitored, why, and who can view this data. A lack of clear answers to these questions can result in a sense of distrust. From the student perspective, if all of the data processing prior to analytic presentation is "black-boxed," it may seem as though analytics are for the benefit of the institution, administration or course designers, and not the students themselves (Slade & Prinsloo, 2013). For mainstream adoption to become a reality (Horizon Report, 2014), there needs to be an established level of transparency and trust between the development and use of analytics.

Interpretation: The Challenge of Priorities. The challenge of priorities reflects the varying decisions learners make in allocating their time and focus when presented with a variety of analytics, in addition to their regular course tasks. Students must decide when and how often to consult the analytics, as well as how they utilize the information provided. This may be problematic for students because even if instructional intentions are clear, learning activities are often complex, and not all useful learning activities are meaningfully captured by analytics. Thus, there is a danger that learners will prioritize the optimization of those things which are reflected by the analytics regardless of their actual importance for learning (Clow, 2013; Verbert et al., 2014). This, in turn, can take away focus from monitoring activities that are valuable for learning but not included in the analytics. Additionally, students may become frustrated if certain aspects of their learning activities are not represented by the analytics—they may develop the perception that these efforts are not valued. Importantly, there is a tension in that learning is a process (Kolb, 1984), whereas analytic metrics often represent the outcomes of this process, overlooking the intermediary steps along the way (Clow, 2013).

Interpretation: The Challenge of Individuality. A challenge students face in interpreting analytics is how to determine what their individual analytics mean in comparison to their understanding of course expectations. This standard is often the class average or a specific target determined by the instructor. The difficulty with this comes from students trying to adhere to a pre-determined standard, rather than focusing on learning. For example, in an online discussion, quality of posts is often more important than quantity, yet the standards for analytics are often set to display class average in quantity form. Students are drawn to this as a standard, to the potential detriment of the larger learning goal of quality correspondence. Additionally, each learner has his or her own starting point, pacing, as well as strengths and weaknesses that may shape how he or she engages with a learning activity. If the analytics determine what aspects of their activity students monitor, they are deprived of the opportunity to set their own pacing and explore their strengths and weaknesses. Analytics need to provide the opportunity for multiple paths for success, otherwise individuality is lost and student decision-making is replaced with compliance to the metric. This deprives students of their role as agents in their developmental and temporal learning trajectories (Slade & Prinsloo, 2013).

Learner Decision-Making.

Decision-Making: The Challenge of Possible Options. Learning analytics are produced and presented to the learner under the assumption that some action could take place as a result of the information provided (Clow, 2013; Ferguson, 2012). Taking action is a key challenge as it addresses the needs of the learner to first feel empowered to make a choice among a variety of options, and second, to use analytics to help inform the decision-making process. Deciding among possible options requires both the flexibility to have multiple choices as well as being given analytics that can help to inform students and support these decisions. Just as analytics need to represent the variety of actions students take in a learning environment, so too must there be choice in their application. Some authors warn that if we are not careful, learning analytics could actually disempower learners by encouraging them to conform to the feedback that analytics provide (boyd, & Crawford 2011; Buckingham Shum, & Ferguson, 2012). Kruse and Pongsajapan go so far as to say that learning analytics can "perpetuate a culture of students as passive subjects—the targets of a flow of information—rather than as self-reflective learners given the cognitive tools to evaluate their own learning processes" (2012, p. 2).

Decision-Making: The Challenge of Enacting Change. Beyond the challenge of being able to make decisions, students are also faced with determining how and when to make changes. Even if the analytics are understood, it may be difficult to link their diagnostic interpretation to actionable steps for improvement. The analytic feedback provided is limited; therefore, students may not know what to do, how or when to do it. Students need guidance in terms of prescriptive steps where analytics provide descriptive representation of past activity.

Decision-Making: The Challenge of Dependency. While the challenges previously outlined are presented from the student perspective, the challenge of dependency comes from a larger concern *for* students. Specifically, students may become dependent on analytics if they are not given the autonomy to interpret the analytics for themselves, as well as the opportunity to make their own decisions. The optimal use of student-generated data is to make better-informed choices (Oblinger, 2012). However, if students are alerted every time they make a small error or misstep, they are missing the opportunity to learn to identify these mistakes for themselves (c.f. Mathan & Koedinger, 2003). Finding a balance will be important as too much guidance may result in students no longer learning to make their own decisions; too little guidance may result in students struggling to determine what their options could be. The goal for students is to learn from the analytics (Booth, 2012), so that they can transfer these skills to new learning activities.

Building a Model for Student Use of Learning Analytics

In the previous section we outlined seven challenges for learners' use of analytics. In this section we consider how these challenges might be addressed by building on Clow's (2013) idea that productive learning analytics use should be considered as part of a cycle of reflective practice (Schön, 1983). We begin by briefly providing an overview of the educational idea of self-regulated learning as a starting point for conceptualizing a reflective cycle of learning analytics use. We then present the *Student Tuning Model* as a specific representation of a reflective practice in which students engage with analytics, interpret, reflect and take action.

The Reflective Practice Cycle in the Educational Literature on Self-Regulated Learning

The idea of learning analytics use as a part of a reflective practice cycle is similar in many ways to the notion of self-regulated learning (SRL). One of the goals of presenting students with analytics is to provide them with a tool for metacognitively monitoring their learning behavior, so that they may use this information to take actions to self-regulate (Roll & Winne, 2015). Numerous theories of SRL describe the process of self-regulation in different ways (for extensive reviews, see Boekaerts, Pintrich, & Zeidner, 2000; Zimmerman & Schunk, 2001); however, all share several premises in common. First, SRL is considered to be cyclical in nature and activity can begin at any point in the cycle. Second, SRL involves three phases of activity: a preparatory/preliminary phase, a performance/task completion phase, and an appraisal/adaption through metacognition phase (Miller, 2015). Importantly, in the SRL cycle learners are conceptualized as planning how to learn and act based on *their* individual goals. If these goals are not met or they find the work is too effortful, they may change tack. Thus goals are considered to be established by learners, rather than external mandated, though certainly there are external factors that may influence them.

Students often engage in SRL, but it is not always a productive process (Winne, 1995). Obstacles to productive SRL arise when learners neglect or ineffectively monitor their progress towards their goals, use inappropriate standards for monitoring, or are unable to change strategies (exercise metacognitive control) because they don't know about or are unskilled in other tactics for learning (Winne, 2011). Learning analytics has the potential to help address many of these challenges by supporting student monitoring as well as setting a context for what to monitor. Beyond this, however, students often still need help in determining what to change.

Due to the multiple ways in which learning analytics data can provide feedback to inform the SRL activities of monitoring and appraisal, there has been a recent surge of research exploring potential connections between SRL and learning analytics (Roll & Winne, 2015). However, uncertainties remain as to when, how and what to introduce as analytics for student monitoring to facilitate learning. Interesting questions are emerging around how students make choices in the learning systems (Cutumisu, Blair, Chin, & Schwartz, 2015; Colthrope, Zimbardi, Ainscough, & Anderson, 2015), the context for these choices (Segedy, Kinnebrew, and Biswas, 2015; Nussbaumer, Hillemann, Gütl, & Albert, 2015) and how these will influence students' use of learning analytics in the future. Addressing the questions of how to best support students in SRL-oriented learning analytics use first requires a conceptualization of the process by which students use data to self-regulate. To address this, we introduce the *Student Tuning Model* as a specific vision of learning analytics use in a cycle of self-reflective and self-regulated learning.

The Student Tuning Model

Drawing on the self-regulated learning literature discussed above, we now describe a specific model of learning-analytics-informed reflective practice. The *Student Tuning Model* is a continual cycle in which students plan, monitor and adjust their learning activities (and their understanding of the learning activities) as they engage in the educational environment with learning analytics (Figure 1). The four elements of the model are laid out below. In the remaining sections of the paper we will turn to the question of how this model of analytics use can be supported through well designed learning analytics implementations.

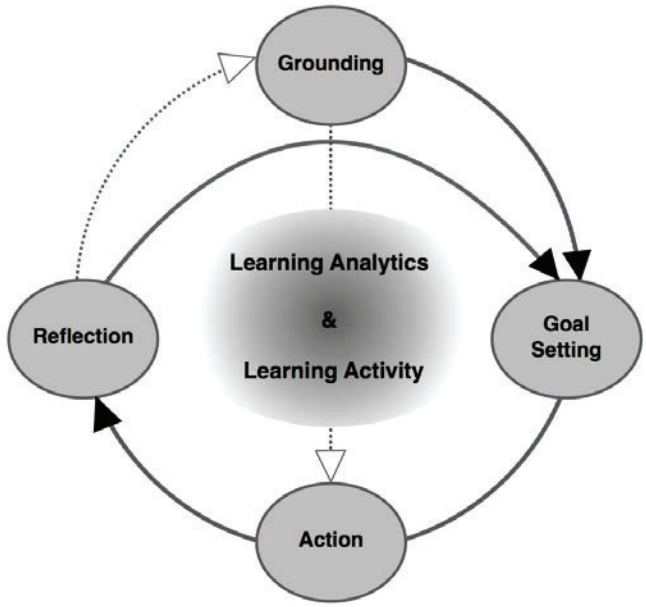


Figure 1. The Student Tuning Model. This figure illustrates the relationship between the four elements of the model.

Grounding. While students may enter the cycle at any point, conceptually the tuning process begins with *grounding*, a process in which students grapple with the general question of what information the analytics provide and how it relates to their specific educational context. Features of the learning activities involved, including their perceived purpose and any expectations for the process or outcomes of engaging in them, may all influence the extent to which particular analytics are seen as useful and relevant (or, conversely, as extraneous and worthless). For students to engage with analytics they need to make sense of the roles the analytics are being introduced to play in their learning environment. For example, analytics that are introduced as a personal reflection tool, part of the course assessment, or to provide recommendation advice, will each be related to differently by students. Later we will discuss how to purposefully support the process of grounding; however, we note here that this activity of sense-making and evaluation occurs with or without intentional action by the instructor. Importantly, beyond intuiting the course-driven context for the analytics, students may also bring their individual contexts to bear. This personalization can take the form of accepting or rejecting (or placing relative emphasis on) particular learning goals, expectations, and analytics or bringing in entirely new elements outside the formal course structure. Such adaptation of the given goals is an example of how students blend normative descriptions of class activities with their own perceptions and values (Pintrich, 2003). In short, grounding defines the overarching relationship of the student to the analytics in a particular learning context. While we have depicted grounding as a distinct element in the model shown in Figure 1, it could also be thought of as a distributed cloud that continuously operates in the background during all other elements of the cycle since it refers to the links students are continuously making between the analytics and their expectations for a learning environment.

Goal-setting. *Goal-setting* deals with the planning of specific proximate objectives and actions for reaching them relative to the larger framework established through grounding. Specifically, learners consider what they want to achieve in reference to what they believe is expected, setting one or more goals accordingly. For example, through grounding, a student might come to recognize that the overall instructional purpose of an ongoing collaborative GoogleDocs writing activity is to synthesize key course themes, of which two in particular interest her. She can then connect this to the fact that the provided analytics show her the frequency and size of her contribution to each theme (e.g. McNely et al., 2012), Specific goals emanating from this grounding could be to contribute regularly to these two themes or perhaps to become the most substantial contributor in the class to them. Once set, these goals drive how students interact with educational materials and activities, and the feedback the analytics provides becomes an important moderator for students to monitor and assess their progress towards their goals, as well as evaluate when the goals themselves need to be updated or revised (Locke & Latham, 2006).

Goals are important in self-regulated learning and analytics use because they can motivate learners to put greater efforts and also incite self-monitoring of their achievement. Self-set goals especially lead to higher self-efficacy which, influences the amount of effort learners make and their commitment to fulfill their learning challenges (Zimmerman, Bandura, & Martinez-Pons, 1992). Finally, goals are useful because they set a standard to which actual activity (as reflected in the analytics) can be compared. While goals logically precede action and reflection in the cycle, students may not always set these prior to engaging in a learning activity or system. Other possibilities include that goals emerge as a by-product of action in the system or as a result of reflection on it. It is also possible that students never become intentional about their learning activities, impoverishing the potential for analytics to support them.

Action. Action is a key element in the Student Tuning Model; *action* is when learners engage in behaviors to realize their goals. These behaviors also generate the data from which the analytics will be created, though it is important to recognize that not all behaviors, nor all qualities of these behaviors will be reflected in the analytics. It is in this stage that the actual tuning activity occurs; that is, that the learner attempts to adjust their learning behaviors to better fulfill their learning purpose. This activity is conceptualized as tuning because it involves regular, relatively small changes (ideally improvements) in response to feedback. The idea of tuning stands in contrast to models in which analytics lead to a radical change, or paradigmatic shift in students' learning activities. There is an interesting asymmetry in that while the tuning action can be considered the heart of the self-regulation cycle because it is where change takes place, the other elements supporting this change (planning, monitoring, evaluation) have been more heavily theorized.

Reflection. *Reflection* builds on the three other phases as students look back at the actions they took (reflected in the analytics) in comparison to the goals they set, all while grounding this in their perceptions and perspectives on the course expectations. This is the key phase in which the interpretation of the analytics and decisions about what changes should be made occur. Specifically, learners may make decisions about any of the other phases: a) setting an intention to change their actions to better meet their goals; b) changing specific goals based on their actions; or c) changing their overarching grounding in how they conceptualize the learning activity. Reflection has long been thought of as an essential part of constructing one's understanding (Schön, 1983); in turn, as one's understanding develops, reflection can also be used more effectively to support learning (McAlpine, & Weston, 2000). Traditionally, reflection has problematically often depended on the learner's own faulty and incomplete recollections (Veenman, 2013). Learning analytics offer an important advantage to this activity by providing data that can serve as an object of reflection. Data can provide a more precise account of the learning process on which to reflect, though this account is neither complete nor neutral. The information provided by analytics can also be considered as a kind of formative feedback. While assessment in education is often negatively associated with (high-stakes) summative testing, when done in a formative way as part of self-regulated learning, it can be a powerful tool to support students in better understanding and taking control of their learning processes. This is an important point as it not only addresses the role of students' intentionality in the system, but also the fact that tuning is a continual process that students continue to engage with as they progress towards their goals.

Supporting the Student Tuning Model: Methods for Developing and Validating a Framework for Pedagogical Design

Above we outlined the Student Tuning Model as a descriptive representation of how students can productively engage with analytics. We now move to describe the methods by which this model was used to generate and test a prescriptive framework for designing the use of analytics in a learning environment. This is a first step to bridge the gap in the literature between *design of analytics* themselves and *designing for analytics use*.

Prelude: Conceptualizing Learning Analytics Implementations, Not Interventions

The primary term that has been used to describe how learning analytics are introduced into the learning environment is “intervention.” We ourselves have used this term previously and found it useful to distinguish between the design of learning analytics tools and the design of the activities in which they are used. While this distinction between the design of technical artifact and social processes is an important one, we are uncomfortable with the language of “intervention” because of several undesirable connotations. Specifically, the term intervention implies learning analytics use as an interruption to regular learning practices at a specific point in time to address a problematic situation (Wise & Vytasek, in review). In contrast, we view learning analytics use as a productive and ongoing part of the regular adjustment of learning practices. For this reason, we use the term “learning analytics implementation” to describe how learning analytics are introduced into the learning environment as we feel it better represents the integration of the learning activity, analytics and student tuning process that we hope to facilitate.

Generating the Align Design Framework for Learning Analytics Implementations

The Align Design Framework was developed as a tool to support the conceptual process of how students engage with analytics described above. The framework is presented as four interconnected principles for pedagogical practice that can support students’ analytics use. Our selection of theoretical ideas and our derivation of guidance based on these ideas was grounded in our understanding of the Student Tuning Model, related research, established critiques of analytics usage outlined previously, and our collective experience as educators and students. This method follows a tradition in education and human-computer interaction of drawing on theory and past experience to generate design guidance that is then tested in practice (Design-Based Research Collective, 2003; Stolterman & Wiberg, 2010). The framework is presented in the Results section below; rather than describe each principle in purely abstract terms, we have combined the presentation of the conceptual design framework with examples of its instantiation in practice. This instantiation described also served as the context for the initial validation of the framework; validation evidence for the principles is described at the end of each section.

Methods for Initial Validation of the Framework

In the initial validation of the Align Design Framework we asked the following question:

How did the implementation of the principles of the Align Design Framework support and/or hinder engagement in a tuning process by graduate students in a seminar course provided with analytics about their online discussion activity?

The validation study was conducted on a small implementation of the E-Listening Analytics Suite in a semester-long graduate-level class on educational technology into which nine doctoral students enrolled. The class met face-to-face once a week for thirteen weeks, with online discussions conducted between sessions for ten of the thirteen weeks. The E-Listening Suite provided analytics about students’ activity in the online discussions, focusing primarily on their attention to others’ messages. Analytics were provided both embedded into a (specially designed) online discussion forum tool and extracted from it and sent to a dedicated analytics space that also included a reflective journal tool. An explanation of how each of the four principles of the Align Design Framework were instantiated in this implementation if the E-Listening Suite into this particular learning context is given in the Results section as each principle is described. For details about the specific analytics provided in the E-Listening Suite and students' reactions to them see Wise, Zhao and Hausknecht (2014).

At the end of the term, all of the students in the class and the course instructor were interviewed by research assistants about their experiences using the learning analytics as part of the online discussion activity. The hour-long semi-structured interviews were organized around four main topics: 1) questions about how the student understood the purpose of the online discussion as well as their participation in it; 2) reaction to and use of the analytics embedded in the discussion tool; 3) reaction to and use of the specific metrics provided as extracted analytics; 4) experience using the reflective journal with the analytics. Transcripts and text of the reflective journal kept by each student were analyzed qualitatively using an inductive approach (Thomas, 2006). Three researchers first marked as relevant text any comments addressing aspects of analytics use. From there, an iterative coding process of identifying repeating ideas, exchanging the coded texts, and dialoguing about the emerging interpretations was used to develop higher-level themes (Auerbach & Silverstein, 2003). All themes were supported by multiple quotes coming from more than one student. The final themes were then organized and examined in relation to the four principles of the Align Design Framework. Because the analytics implementation that students experienced was based on the framework, all themes addressed at least one principle.

Results - The Align Design Framework and Initial Validation Evidence

The Align Design Framework (Figure 2) consists of four interconnected principles for pedagogical practice (Integration, Agency, Reference Frame, and Dialogue/Audience) to support students' analytics use. The following sections introduce each of the principles in three parts. First, a conceptual description of the principle is given, including a theoretical rationale for how it supports the Student Tuning Model and addresses the identified challenges associated with students' use of learning analytics. Second, a practical account is provided for how the principle was instantiated in the validation study. Finally, findings from the study are presented to demonstrate how the principles supported and/or hindered students' engagement in the tuning process in this context.

Principle 1: Integration

The central goal of the Align Design Framework is to provide an intentional context for the human activity in which analytic tools, data, and reports are taken up and used. Thus the first principle in the framework is the *Integration* of the analytics into the overarching learning activities taking place. The goal of Integration is for the instructor or designer to position analytics use as an integral element in the learning process tied to their goals, expectations and planned learning process.

Integration can be thought of both conceptually and practically. Conceptually, integration involves helping students to understand the goals of the learning activities for which analytics are provided, recognize what is considered as productive engagement towards these goals, and finally make the connection to how the analytics provide indicators of the ways in which their actual activity does or does not match these expectations. In this way, conceptual Integration can help to support the student tuning processes of grounding, goal-setting, action and reflection in meeting the interpretive challenges of context and priorities discussed earlier. Importantly, not all of the analytics provided by a system may align equally well with the goals of a particular learning activity or even an entire course; thus part of designing for Integration is figuring out which analytics should be focused on in a given situation, and which should be de-emphasized. Further, it may be necessary to point out valued qualities of activity that are not represented by the analytics. Thus the overall goal of conceptual Integration is to create a foundation of the qualities of productive participation in a learning activity which can act as a local context to make sense of and act on the data. Thus, the same analytics suite can be tailored to meet different needs across diverse learning contexts.

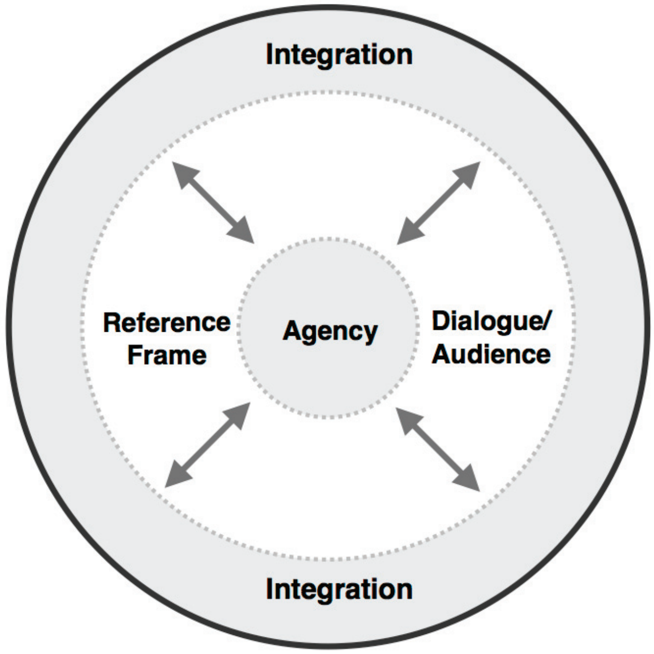


Figure 2. The Align Design Framework. This figure illustrates the relationship between the four principles of the design framework.

Practically, Integration is also about how the use of learning analytics is incorporated into the activity flow of the educational environment so that their use is a regular part of the learning process, rather than ancillary to it. This incorporation might include dedicated upfront time to establish a shared understanding with students about the conceptual connections between purpose, process and analytics (or perhaps to co-construct this understanding with the students), as well as to explain to some extent how the analytics are generated from the data. The depth to which this latter issue is treated will depend on many factors, including the complexity of the calculations and analytical sophistication of the students; however, some attention is needed to address the interpretive challenge of trust. In addition to these upfront concerns, practical integration also means helping students set a rhythm for the goal-setting/action/reflection tuning cycle. For experienced learners this may come naturally, but for many students, guidance about when and how often to consult the analytics, assess progress towards goals and make adjustments to their learning can provide useful structure (and restraint) for their analytics use.

The frequency with which the analytics are provided or accessed as well as the schedule for reflective activity will vary depending on the context. The goal is to create a specific timing for cyclical review so that students are not overwhelmed by or become overly reliant or fixated on the analytics (Buckingham Shum & Ferguson, 2012). Analytical feedback needs to be provided quickly enough to impact practice (Buckingham Shum & Deakin Crick, 2012) but also on a scale for which the analytics make sense to examine in a particular context. Especially for analytics that track larger scale constructs, the time-frame over which the data is examined can dramatically affect the results (Zeini, Göhnert, Krempel, & Hoppe, 2012). Guiding students around the structure and timing of this is a crucial part as it sets the stage for students to learn how they can interpret and use the information provided through analytics to monitor their progress towards their goals.

Implementing Integration. In the initial implementation of the Align Design Framework, Integration was enacted in several ways. Conceptually the instructor mapped the goal of the online discussion activity (to build individual and collective understandings of the course material through dialogue) to her expectations for productive participation in it (guidelines for the quantity, quality, and timing for making posts as well as broad, deep, integrated and reflective attention to the posts of others). The analytics were introduced in this context; the instructor described how information embedded in the discussion interface could support particular elements of participation and the metrics extracted from it mapped to the expectations (e.g. the analytic of "percent of posts read" was an index for the breadth of attention to others' posts while the interface visually showed the distribution of these posts throughout the discussion, indicating if this attention was integrated or scattered (see Wise, Zhao, & Hausknecht, 2014 for details). The instructor also highlighted for students that while analytics about posts' contents were still in development, the quality of contribution was nonetheless an important element of the expected participation. It was also made clear that students had flexibility in the interpretation of their analytics, which was encouraged through the reflection activity, described in further detail under the principle of Agency.

Practically, Integration was enacted through clearly articulating the linked goals, expectations and analytics described above in initial discussion / analytics guideline documents which were presented and explained in-class; students also had the opportunity to ask any questions they had about the analytics, their collection and calculation, and their use in the class. The analytics were explicitly presented as tools for reflection to understand and make changes in one's discussion participation, rather than a part of the course evaluation. The course was then structured so that students were guided through a weekly reflection journal activity to set goals for their discussion participation and consider their progress towards these based on their self-reflection and the analytics. This set-up a week-long rhythm of goal-setting/action/reflection.

Validating Integration. Initial evidence suggests that the elements of Integration described above were useful in supporting students in the tuning cycle in several ways. At a basic level, students seemed to be supported in grounding their thinking about the analytics in the expectations outlined for the discussion activity. One student shared that the analytics were "trying to tell you about meeting a goal that we had and set up… so I get that." Evidence of grounding was also seen in the way students discussed the guidelines for discussion participation and the analytics as one unified concept which informed their understanding of the expectations for participation. As one student explained:

> [Instructor Name] was proactive in that in setting up those criteria. She did a lot of work to make it work to introduce rules of engagement at the beginning—that helps. I think it was very good to have scope outlined very clearly at the beginning and I think that did help actually…The other thing was just setting reasonable parameters of what is appropriate and what makes sense and what you are trying to do. The general guidance worked really well, they were well thought out and they definitely informed the scope of what we should be doing in the discussion forum.

Beyond establishing the scope of productive activity, many students indicated that they actively thought about the guidelines and analytics while participating in the discussions and used these to monitor their activity. For example, students were given the guideline to keep their posts relatively short to contribute to an ongoing interactive dialog and the analytic of post length was available to monitor this. One student shared how they used this information while participating in the discussions: "If I felt my word count was too big, then I wouldn't hit the reply, I would actually think about it a little more…". Another student mentioned that they thought about the range of participation analytics when deciding when to log-in to the discussions: "I was working specifically on trying to get in earlier regularly." One student went so far as to suggest that the integration could be made even tighter by linking the guidelines directly to the discussion so people would not ignore or forget the guidance while engaging in the forum.

Students reflected that they saw these guidelines and analytic metrics working together as flexible to meet the needs of the discussion. One student shared, “You know, you don’t want to say to people, ‘Never post beyond 300 words’ ‘cause sometimes there is a place to do that, so I think it is good to have it kind of open.” The analytics and guidelines were “setting reasonable parameters of what is appropriate and what makes sense and what you are trying to do. The general guidance worked really well.” These guidelines were established by the instructor who students thought “did a lot of work to make it work to introduce rules of engagement...that helps”.

While the students were made aware of the analytics as well as the potential benefits of their use, some still found it challenging to decide when to consult the analytics and how to make them part of a self-reflective tuning cycle. One student shared, “It is interesting because I did have access to those analytics persistently throughout a week, [but] I would tend to forget about it in the course of being involved in the discussion forums.” This quote highlights that linking engagement in the activity and analytics use requires more than simply making this opportunity available. Similar to the more general process of self-regulation, it could be that engagement in the student tuning cycle has multiple individual factors that go into how, when and to what extent a student chooses to engage. There was also evidence to suggest a challenge in that some students lacked trust in the analytics, which may have dampened overall use. As part of the introduction to the analytics, students were given an explanation of how their activity data was being tracked and presented, however some students still doubled the validity of their analytics. As one student said, “I did try [to use the analytics] but at the same time when there is that little bit of doubt and you think, ‘I don’t know if this is really measuring that’ it seems like it might not be that accurate and then I – I didn’t withdraw from it, but maybe I put it down in priority a little bit” indicating his partial dismissal of the analytics. It seems that greater attention to garnering student buy-in to the analytics may be a key element in encouraging their uptake.

In summary, evidence suggested that the implementation of Integration in this context supported students in composing a unified understanding of the learning activity expectations and analytics. Students used this understanding to guide and monitor their activity in a flexible way; however, challenges were found in student trust of the analytics and their ability to make them part of their regular patterns of activity.

Principle 2: Agency

The purpose of Integration was to establish an instructor-driven frame for the learning activity and analytics use; however, it is also important for students to take ownership of their learning process and be proactively engaged in their own learning to be successful (Zimmerman & Schunk, 2013). In fact, the possibility to support learners in actively taking charge of managing their own learning process is one of the key attractions of learning analytics (Govaerts, Verbert, Klerkx, & Duval, 2010). For these reasons, the principle of Agency is at the center of the framework, directed at giving students the opportunity to engage with analytics as a tool to inform their actions, as opposed to analytics being something with which students must comply. Agency provides a counterbalance to Integration, promoting a degree of equilibrium between instructors’ and learners’ authority in setting intentions for activity and monitoring the results. While, by definition, Agency is not something that can be imposed through design, learning and analytic environments can be constructed in a way that creates space for and encourages it.

There are three important conceptual dimensions to consider when designing to support Agency. The first dimension relates to creating an environment that allows for individualization of learning goals. By making personalized goal-setting an explicit and structured part of the learning activity, learners are asked to be purposeful in thinking about the stated objectives of the activity, evaluate their own strengths and weaknesses, and set specific and proximal targets to work towards. The process of goal-setting should be tied to and follow from the introduction of the learning activity's purpose as described above under Integration. In this way, Agency supports students in balancing these external goals and expectations with their own internal ones. Such balance is important because each student has a different starting place and skill-set that they bring to their learning, so they may need focus on different aspects of the learning task that require more attention than others. In other words, there is a need for multiple profiles of productive activity and improvement, rather than a single goal and path to which all students must aspire. These individual goals can then provide a personalized context for the sense-making of analytics, the second dimension of agency.

This next dimension of Agency relates to flexibility of interpretation. While the instructor generally determines the selection of analytics as well as their formal role in the course, they can also make efforts to create space for students to have some authority in interpreting their meaning. Providing this space is important because students have access to information instructors do not, which can contextualize the analytic data within students' personal experiences; for example, a student will know whether they spent especially long working with particular content because they were very interested in it, or very confused. These aspects of interpretation are an important part in making sense of the feedback the analytics provide and how it is relevant to students' learning. Instructors can cultivate recognition that students' interpretations are appreciated and respected, so that students begin to value the feedback they provide for themselves through self-evaluation. For example, instructors can refrain from providing ready-made interpretations of data, instead modelling different ways the information can be thought about and inform the learning process.

The third and final conceptual dimension of Agency surrounds student decision-making based on the analytics. Here instructors must work to strike a balance so that students can decide for themselves what actions to take, but also provide enough guidance so that they are not paralyzed with uncertainty about how to take action. There is tension between too much guidance which could lead to student over-dependence and too little guidance which doesn't give students the opportunity to take data-informed action in a timely way. Thus instructors need to create the space for students to make decisions about action, while also providing some guidance about how to do so. This guidance may be offered through prompting questions that remind students to look back at previous goals, consider the analytics, and think about where and how changes can be made. Additional forms of responsive guidance are discussed under the principle of Dialogue.

In application, one powerful way to support Agency is to provide a student-owned space for receiving, working with, and documenting thinking around the analytics. Providing a dedicated place for this activity to occur establishes it as a valued activity in the learning process. Through documenting their thinking around the analytics, students can explore the identification of their own goals, interpretation of analytics and how these can inform decisions about future action. For example, a popular technology for creating an individually-owned reflective space is private or publicly shared blogs (Ferguson, Buckingham Shum, & Deakin Crick, 2011).

Implementing Agency. In this implementation, students were each given a personal and private wiki space in the form of a journal in which to set goals and reflect. This is also where their extracted analytics appeared to them, provided on a weekly basis. The instructor encouraged students to use the space to reflect on their progress each week as well as update goals with regards to their planned activities. Students were provided with a small amount of class time for doing this and the instructor served as an audience for the journal providing occasional comments. By dedicating (technical and temporal) course resources to the reflective wiki journals, the instructor aimed to instil a sense of value in students' reflective practice and established the expectation that students would come to their own understanding of the analytics as well as what role it would play in their learning. Students thus had the first word (literally) in composing the meaning of the analytics. They could also ask questions and receive guidance if they were struggling to make sense of the metrics or change their learning behaviors based on them. In a limited number of cases, the instructor drew students' attention to metrics they may not have focused on or suggested an alternate interpretation to the one put forth by the student, but the core of the activity remained student-driven, centered on their goal-setting and what they valued as part of the learning activity.

Validating Agency. Initial evidence suggests that the implementation of Agency facilitated student-driven goal-setting and cyclical reflection. Students themselves recognized the role of the reflective journal in helping them to take ownership and control of the process, pointing out "it definitely improved my awareness of what I was doing in the discussion. So when I would go into the discussions, I definitely could see myself being more self-regulating …" Specifically, the development of this process-level insight was seen in the goals that student set and the way they chose to focus on different components of the analytics. As one student explained, "There were some [metrics] that I put more weight on than others." For example, one student shared that as a result of reflection her goal was to "go into the discussions when I had time to really actually think about things as opposed to just, you know, read them, check, read them, check." With previous reflections inspiring plans for future changes she noted, "I would typically read the previous weeks' goals and the feedback from the instructor again and just make sure that I did sort of cover those bases."

Students also demonstrated ownership by setting the criteria for when a goal was completed. As one student explained,

> Generally, I tried to shift the goals. I did have some things that I would come back on that I felt that I wanted to try this again or think that I've actually – I don't think I received enough feedback within the forum itself to gauge if that had shifted. So I felt like I could try for two weeks or maybe three. Yeah, but each time trying something different was an interesting approach.

Through this process, students felt capable of and invested in achieving their goals. One student proudly shared, "Pretty much every week where I did set those goals I was pretty much able to make those changes… it was a good opportunity to think about that and actually plan for it." Further evidence that the implementation of Agency supported student ownership of the tuning cycle was found in goal-setting activities that evolved beyond course guidelines to represent actions students thought would be personally valuable to their learning. One student commented that, "One of my weekly goals sort of midway through the course was to try and organize all of the postings that kind of resonated with me." This was not a strategy introduced during the course but rather one that came from the learner as part of the goal-setting and reflection process.

Going beyond goal-setting, students showed evidence of establishing their own interpretation of the analytics provided to make decisions about what changes they did and did not want to make. Firstly, students brought information from their experience participating in the analytics to bear in their interpretation. For example, in looking back at their analytics, one student noted in their reflective journal, "yeah, things looked low for the week, but this was a busy week, it will be better next week." Secondly, students recognized that different profiles of activity were not necessarily better or worse, but just different. For example, one student shared that "[it] was interesting to compare myself to others because, again, [I like to express] myself in short paragraphs instead of long things. I like short and condensed and meaningful ….we [are] all different, [it is] not that something is bad or good, it is just different things." Thirdly, students were aware that they could select their own path for change, commentating "I think based on the numbers you could have… you could go in a direction or choose the direction how you want to go." One example of this is a student who described,

> I could see I wasn't in the discussion for as many days as I would have liked, so I would make an effort to go in, even if was only briefly, on the weekend so I could sort of, you know, be a part of the discussion development more.

Of note in this quote is not only the student's decision to make a purposeful change to their discussion activity based on the analytics, but that this change was tied to a motivation about the discussion's purpose and not simply the numbers themselves. Finally, a different student pointed out that

> …knowing that there is a metric that would tell me how much I really spent time on was good because it made it feel okay …I won't click on every single thing, but at least I'm going to make sure I try and spend time alone with what I do.

This illustrates how students used the analytic as a flexible tool, as opposed to the analytics rigidly determining their actions.

Despite the instructors' efforts to provide space for Agency, not all students took up the opportunity to take ownership in goal setting and the interpretation of the analytics. For example, one student described "she always asks us to write a goal…But some days I'm too busy to just read and post something and I forgot what I wrote about my goal." For other students the problem lay not with the goals but their ability to take action to meet them: "I thought it was useful to reflect, but I found it frustrating to reflect and set goals and then not necessarily be able to actually meet them." Finally, some students simply chose not to engage with the process in a meaningful way: "I don't actually think my participation changed very much. I don't think writing about that reflection piece contributed to any changes that I made in my participation."

In summary, the validation evidence suggests that the implementation of Agency in this context supported many students in taking ownership of their analytics use by setting individual goals, actively interpreting their analytics, and making personal decisions about changes. While the majority of students demonstrated taking ownership of their analytics use in some way, some did not engage with the analytics in an agentic fashion. This reinforces the point that Agency is not something that can be directly imposed but rather must be indirectly supported and encouraged.

Principle 3: Reference Frame

The initial principles of Integration and Agency create an overarching structure surrounding the activity and promote a core of student ownership; the final two principles target support for the space of negotiation between instructor-driven Integration and student-driven Agency as well as the contextualization of learning analytics use. The first of these principles is Reference Frame. Reference frames are the comparison points which orient students' interpretation of analytics. In general, there are three potential reference frames to which analytics can be compared. The first is the criterion articulated as part of the course expectations, which were described under the principle of Integration. These provide an absolute reference point to which students can compare their own analytics. A second reference frame for comparison is a student's own prior activity as discussed under the principle of Agency. This is a relative reference point since a student's prior activity does not represent an immovable goal, rather one which supports the tracking of improvement over time.

The final reference frame for comparison, and one which has not yet been discussed is the activity of other students. Aggregated information about the performance of others student is often provided in analytic systems and can be powerful in showing a student where they stand in relation to their peers (Govaerts, Verbert, Duval, & Pardo, 2012) but can also have negative consequences for their self-efficacy. For example, low-performing students who may not initially realize how their efforts stack up against others might be motivated by peer comparisons or find this information stressful and intimidating. In addition, by definition a student's peer provides a relative, not absolute, standard for comparison. Thus, if a class average is substantially lower than an instructor's expectations, it may not be an appropriate target to aim for. In other cases, everyone may already be well beyond the bar of what is necessary, making additional exertion to improve a particular metric wasted effort. In particular, measures of the class's central tendency (particularly the average) have several challenges as a reference point, including the potential to be overly influenced by the activity, or inactivity, of certain students and flattening out differences across particular subsets of students. For these reasons, instructors need to consider how they introduce peers as a reference frame so that it supports a productive, not detrimental tuning process.

Some of the issues described above can be addressed through careful design and refinement of analytic tools in how the peer reference frame is presented, for example, processing data to provide aggregate measures for only similar kinds of students or providing aggregate measures of spread (variance) as well as central tendency. However, there is also an important role for pedagogical design to play in terms of helping students to balance the different reference points as well as understand their value and limitations in a specific context. In practice providing reference points can take several different forms depending on situational factors. In some courses external standards for expected activity can be emphasized as a fixed guidepost by which to judge progress. In cases where absolute indicators are harder to provide, the instructor might stress the importance of the personal reference frame, explicitly asking students to monitor the changes in their analytics over time. The overall aim in each of these designs is to help students avoid the simplistic mentality of "more (than other students) is better."

Implementing Reference Frame. The course implementation attempted to balance the use of all three reference frames. As previously mentioned under the principle of Integration, the instructor articulated the course expectations in conjunction with analytics, which became part of the first reference frame. In addition, the peer reference frame was made available via extracted analytics that provided students with the group average for each of the metrics. Finally, as discussed under the principle of Agency, students kept ongoing reflection journals in a wiki with the new analytics being added to (rather than replacing) prior ones. This allowed students to make use of their own prior activity as a reference frame. This was actively encouraged through one of the reflection questions which asked, "How does your participation in this week's discussion compare to previous weeks?"

Validating Reference Frame. As reported under Integration, students used the analytics in conjunction with the course guidelines as a reference frame to guide their behaviour in the course and compared their metrics back to the expectations indicated by the instructor at the start of the course. Students also showed some evidence of using their own prior activity as a reference point for the analytics, for example, one student arrived at a realization that they needed to focus their discussion efforts based on seeing "how the numbers are actually decreasing, meaning that you organize your time more constructive[ly] and saying this is something I just want to do instead of *wow*, I want to do everything." Despite some use of students' own prior activity as a reference frame, this seems to be an element that could be strengthened in future implementation designs.

Finally, despite their under-emphasis in the pedagogical design, the peer reference frame was still a very salient one for students. As one student explained, "You can actually see from the numbers how students are, where they are standing, so that was a good one and then you can compare yourself to others." Many students found this reference point to be a useful one: "I kind of see where I [am] among other participants. Then, also I could see my tendency or my approach to participate in the forum" and noted that the peer average was useful to see, "not that it was good or bad, because some people would come in more or less, but it was interesting to see who was doing what." For these and other students, the peer average was not an intimidating force, but one reference point among several. As one student explained, “I was different than the average quite a bit, but then I don’t think you should always be the average, it’s different for everybody, and people need to work on it.” In contrast, for two students in the implementation the class average as a reference point seemed to have a negative influence. For one student it caused a high level of stress and worry, stating that “[Analytics] is the factor which gave me more negative thought about my own way to participate in the forum… always my score is below others.” For another student, the peer reference frame led to the competitive activity of trying to myopically out-perform others on one particular metric. He shared,

> Yeah, usually [I] say in the reflection journal that I should post more or something because I want to reach the average of something, although I know that not every post [is] very meaningful or very significant, but I want to make that number, you know, psychologically.

In this way, the peer average as a reference frame for the analytics detracted from the goal of the learning activity, which was to make meaningful contributions to the discussion.

In summary, students recognized the use of all three reference frames in their analytics use. First, students compared their analytics with the expectations set out by the instructor in the course guidelines for self-evaluation and to guide future participation. Second, there was some evidence that certain students used their own prior activity as a comparison point, however this is a reference frame that could be further strengthened. Despite efforts to balance the use of the three reference frames available, the activity of other students in the class as represented by the peer average was by far the most common reference point used. All students expressed an awareness (both in their journals and in the interviews) of the class average given in the metrics. Many appreciated it as a reference point, using it as a tool to reflect on their own participation by noticing differences, evaluating what that could mean for themselves, and using that to guide future actions. However, for a small but notable minority, it induced feelings of inadequacy or competitiveness, limiting its usefulness for them as a productive reference frame.

Principle 4: Dialogue / Audience

Similar to Reference Frame, the final principle in the Align Design Framework targets support for negotiation between instructor-driven Integration and student-driven Agency. This principle was originally conceived of as Dialogue, focused on setting up analytics to act as an object of conversation between instructors and students. However, through the implementation of this principle and examination of its enactment, we came to recognize that the value provided was broader than simply the opportunity for the instructor to respond to students about their analytics, but also included the basic creation of addressees for the student goal-setting, action and reflection cycle. In this sense, this principle might be reconceived in terms of Audience (to one’s self, one’s peers and the instructor). Because the question of which conceptualization is most productive remains open, we refer to the principle in the current writing as Dialogue / Audience.

There are several reasons to promote Dialogue / Audience as part of learning analytics use. First, creating an audience for goal-setting and reflection can enhance the gravity and sense of commitment to those processes and thus shape the reflective activity itself. Second, these audiences can contribute to making sense of and acting on the analytics. For example, the input of an instructor or one’s peers can be useful in providing an alternate perspective on the analytics and thus facilitate the process of reflection (Andrusyszyn & Davie, 1997). Instructors, in particular, have the opportunity to examine students’ process, use of analytics and respond as necessary with support to address any confusion or challenges the students encounter, as well as query into particular interpretations or goals. Peers too, as other individuals participating in the activity and receiving analytics based on it, can potentially provide helpful insight and feedback. In these conversations, the analytics themselves can act as a third “voice” in the conversation. This gives all parties a neutral object to which they can usefully refer in conversation (e.g. “I noticed that my/your level of participation differs in comparison to the rest of the class” rather than the groundless “do

I/you need to participate more?"). Finally, a Dialogue / Audience gives students the opportunity to ask for help and the instructor the chance to provide suggestions or strategies if students have identified goals based on their analytics but do not know how to make progress on them. This is a potentially powerful way to address the challenges of possible options and enacting change discussed earlier.

Students also benefit from sharing the interpretation of their own analytics with their instructor or their peers. Such communication both acknowledges their voice in concluding what the analytics mean (as discussed under Agency) but also can contribute to the understanding of the instructor (or peers) who may not be privy to all of the information a particular student brings to bear in interpreting their analytics (e.g. "I had a really difficult time with this part of the assignment," "I tried extra hard this week," "I know I need to share my ideas more, but I don't always feel confident that I have the right idea"). This information can help to contextualize the instructor's own understanding and bridge the gap between student and instructor interpretation of the analytics.

A major challenge in enacting the principle of Dialogue / Audience is the issue of scale. In a small class it is possible for the instructor to interact with all students on a relatively frequent basis, but as the student-to-instructor ratio rises this interaction becomes progressively more difficult, and in the case of massive open online courses, it is simply impossible. Two possible alternatives for fostering Dialogue / Audience around analytics are plausible. First, a tiered system could be employed where teaching assistants or student leaders serve as primary dialogue partners, with questions or concerns elevated to the instructor as needed. Second, as previously mentioned, in some situations it may be viable for students to support each other through partnership or triad models. The concern here comes from a lack of experience on the part of the students and the ability to effectively support each other, thus this approach may only work with learners who are relatively proficient in using analytics to support their learning.

Implementing Dialogue/Audience. In this implementation, Dialogue/Audience was enacted by making the student-owned reflective journal a shared artifact between each student and their instructor. Thus the instructor (but not peers) was targeted as the audience / dialogue partner in this situation. Using the wiki technology described earlier, the instructor could both view each students' reflective journal and add text to the journal in a different color than the one used by the student. Students were aware that the instructor had this access and would be reviewing the journal periodically, adding comments and responses at times. They were also encouraged to note anything related to the analytics that they wanted feedback from in this space.

Validating Dialogue/Audience. Students found the implementation of instructor Dialogue/Audience as useful in several ways. First, it provided a "safe space" for students to express themselves. For example, one student wrote in their journal, "I realize I should maybe share my opinion and not just ask questions, but I find that hard." This also gave the instructor an opportunity to address such concerns in a personal way. As another student noted "Because those comments that the instructor gave to me weren't shared with other people…here is a spot where you have someone just responding to you so that was useful." Second, the instructor's comments supported and encouraged students' discussion participation efforts. For example, one student commented explicitly that "the responses were encouraging / supportive" and "[the instructor] is a very collegial and supportive person and that comes across in the feedback." For some students, instructor comments boosted their confidence that they were putting their efforts in the right direction. For example, one student noted, "I do think that the interaction with the instructor was probably the most important thing and I found it reassuring, especially being new to that environment, that I was on track." For other students, the instructors' comments gave them a needed push to make changes to their online discussion participation: "I think (the feedback) is very useful because I got a lot of good or positive results from the instructor... So that pushed me from the safe place because I have confidence to express my idea."

Third, students found the instructor's feedback useful in giving guidance to help them meet their participation goals. For example, one student explained that it was "constructive in terms of, you know, providing some feedback on how to maybe do things differently or to think about doing things in a different way, so it kind of planted some ideas on how to make changes—little tweaks." Another student explained at more length:

> I have ideas of how I should participate for the forum, so that's ideal and sometimes makes me crazy because I should do this, I should do this, but when she said not necessary to do that, why don't you try this way, then I kind of—I thought, 'Oh, okay, I don't have to do this? I could also do this way?' And then that has kind of helped me to release some stress [chuckle].

Finally, the instructors' presence in the shared wiki space during reflection was seen as establishing accountability for the goals students set for themselves. Students mentioned that they tried to engage more because of the awareness of the instructor's observation. One student explained that

> …[the instructor] would read our goals, so I think she maybe plays a role, like as our monitor, so I know my goal to the instructor. So every week I have to did my best to meet the goal… It is just like my own goal and the instructor we go hand in hand.

Even though goals were not evaluated, the process of articulating a goal, even to oneself, provided a sense of accountability that students strove to meet. One student shared

> I will say that in the reflection I need to do much better next week, so I want to make my words. I just don't want to put the words there and without the improvement in participation. So I think that kind of give me some pressure, give me some push to do that.

In summary, the enactment of Dialogue /Audience through making the reflective journals open to and commentable by the instructor was found to be valuable in establishing accountability, supporting student confidence, and providing feedback about how to enact change. As peers were not engaged as a target of Dialogue /Audience in this implementation, no conclusions about their potential role in enacting this principle can be drawn.

Conclusions

This paper has reviewed the challenges for students' learning analytics use identified in the literature, presented the *Student Tuning Model* as a conceptualization of the process by which students use learning analytics as part of a self-regulatory cycle, and proposed the *Align Design Framework* as a set of four interconnected principles to support such use. The framework was enacted in an initial implementation with validation evidence provided about how the four principles of Integration, Agency, Reference Frame, and Dialogue/Audience which comprise the framework support and/or hinder students' engagement in the tuning process. Below we discuss the implications of this work, as well as limitations and directions for future research.

The initial validation study provided substantial evidence that the principle of Integration could be used to stimulate students to construct a unified understanding of the analytics and learning activity expectations, which was useful in guiding the tuning cycle. This understanding addressed the challenges of context and priority in analytic interpretation discussed earlier. Similarly, there was also evidence that the principle of Agency could support (the majority) of students in taking personal ownership of their analytic goal-setting and interpretation, thereby addressing the challenges of individuality and dependency.

The combination of Integration and Agency seemed to support most students in using the analytics in a flexible and personally relevant way; however, supporting student use of analytic requires continued effort to maintain equilibrium between instructional goals and student ownership. In particular, beyond the tension of students having the freedom to own the process yet still align with the overarching parameters of the learning activity, there is the issue of providing sufficient, but not undue, guidance to students who have difficulties interpreting or figuring out how to act on the analytics. In addition, it remains to address the difficulties in students' trust in the analytics and making their use part of students' regular patterns of activity. One particularly fruitful avenue of investigation may be to take reflection on analytics in the context of activity expectations as a starting point for goal generation. This would provide students with analytics as a primer to think more precisely about the process of self-monitoring to track progress towards certain goals. This might also help address questions around the trust in the analytics since goals would be set in the context of a more realistic understanding of how the analytics reflect activity in the system.

The principles of Reference Frame and Dialogue/Audience were introduced as tools to support the negotiation of the friction between instructors' intentions and student ownership. Despite efforts to emphasize course expectations and one's own prior activity as useful reference frames, students focused on the reference point of their peers' activity. This had positive effects for some in providing a bigger picture of how others were engaging, but negative effects for others which resulted in feelings of inadequacy or competitiveness. The audience of and feedback provided by the instructor was generally found to be a positive element in establishing accountability, supporting student confidence, and providing feedback about how to enact change. Enactment of Dialogue/Audience in situations where individual instructor interaction is not possible and continued efforts to constructively shape and offset the use of peers as the sole reference point are areas for future research. The latter is particularly important in situations where peer activity differs substantially from the instructor's expectations. We note additionally that in larger classes, attention will be needed to identify the appropriate form that a peer reference frame takes. For example, if students in a large class are broken into small groups for discussion, it is not clear if the most useful reference point is the activity of these near peers (whose activity can also be witnessed directly) or the activity of the class as a whole (whose analytics provides a broader, but more detached context).

This was an initial small-scale implementation and validation study, and as such, there are limitations to the generalizability of the findings. As mentioned previously, it is important to explore how the principles in this framework can be applied in large-scale learning contexts with vastly greater student-to-instructor ratios. In addition, the E-Listening Analytics Suite used in this work was robust in providing diverse complementary metrics, which were amenable to flexible interpretations. Further exploration into the application of the principles with different analytic systems will be important in helping to understand the range and diversity of situations to which the framework is applicable. We have currently begun such work applying the framework to design the use of a different suite of discussion forum analytics. Apart from employing different kinds of analytics suites, it is also important to work in the context of different kinds of learning activities and content domains. By nature, online discussions are a pedagogical tool in which personalized student goals and improvement profiles make sense. We believe this holds true in other learning activity contexts and domains as well (including those in which very clear "right" and "wrong" answers can be identified), however this remains to be empirically tested. We expect further support and refinement of the framework to come through time as the principles are taken up by learning analytics practitioners and researchers and investigated through both experimental and design-based research approaches.

In conclusion, this paper has taken initial steps towards developing design knowledge for the "middle space," between data presentation and analytics-driven action, with a focus on the case of students as analytics users. We believe that attention to designing for analytics use marks the start of an exciting new branch of research that shifts focus from big data as an object of enquiry unto itself to the large impact on learning that it can engender.

References

Aguilar, S. J. (2014, March). *Exploring and measuring students' sense-making practices around representations of their academic information*. Paper presented at Learning analytics and Knowledge Conference 2014, Indianapolis, IN, USA.

Andrusyszyn, M. A., & Davie, L. (1997). Facilitating reflection through interactive journal writing in an online graduate course: A qualitative study. *The Journal of Distance Education, 12*(1/2), 103-126.

Arnold, K. E., & Pistilli, M. D. (2012, April). Course signals at Purdue: using learning analytics to increase student success. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 267-270). ACM.

Auerbach, C. F., & Silverstein, L. B. (2003). *Qualitative data: An introduction to coding and analysis*. New York: New York University Press.

Avramides, K., Hunter, J., Oliver, M. & Luckin, R. (2014). A method for teacher inquiry in cross-curricular projects: lessons from a case study. *British Journal of Educational Technology*, 46, 2, 249–264.

Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In *Learning Analytics* (pp. 61-75). Springer New York.

Boekaerts, M., Pintrich, P. & Zeidner, M. (2000). *Handbook of self-regulation*, 417-450. San Diego, CA: Academic Press.

Booth, M. (2012, July/August). Learning analytics: The new black. *EDUCAUSE Review*, 47(4), 52-53

boyd, D., & Crawford, K. (2011, September). Six Provocations for Big Data. Paper presented at A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society, Oxford Internet Institute, UK. Retrieved 12 June, 2012, from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1926431

Brooks, C., Greer, J., & Gutwin, C. (2014). The Data-Assisted Approach to Building Intelligent Technology-Enhanced Learning Environments. In *Learning Analytics* (pp. 123-156). Springer New York.

Buckingham Shum, S, & Deakin Crick, R. (2012). Learning dispositions and transferable competencies: pedagogy, modelling and learning analytics. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 92-101). ACM.

Buckingham Shum, S. (2012). *Our learning analytics are our pedagogy*. Keynote address given at Expanding Horizons 2012. Macquarie University, Australia.

Buckingham Shum, S. & Ferguson, R. (2012). Social Learning Analytics. *Educational Technology & Society*, *15*(3), 3-26.

Buder, J. (2011). Group awareness tools for learning: Current and future directions. *Computers in Human Behavior*, *27*(3), 1114-1117.

Butler, D. L., & Winne, P. H. (1995). Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research*, 65(3), 245 -281.

Campbell, J. P., DeBlois, P. B., & Oblinger, D. G. (2007). Academic Analytics: A New Tool for a New Era. *EDUCAUSE Review*, *42*(4), 40.

Campbell, J.P. & Oblinger, D.G. (2007) Academic Analytics, EDUCAUSE Quarterly. October.

Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning*, *4*(5), 318-331.

Clow, D. (2013). An overview of learning analytics. *Teaching in Higher Education*, *18*(6), 683–695. doi:10.1080/13562517.2013.827653

Cuban, L. (1986). *The classroom use of technology since 1920*. New York: Teachers College Press, Columbia University.

Colthrope, K., Zimbardi, K., Ainscough, L., & Anderson, S. (2015). Know Thy Student! Combining Learning Analytics and Critical Reflections to Increase Understanding of Students' Self- Regulated Learning in an Authentic Setting. *Journal of Learning Analytics*, *2*(1), 134–155

Corrin, L., & de Barba, P. (2014). *Exploring students' interpretation of feedback delivered through learning analytics dashboards*. In B. Hegarty, J. McDonald, & S.-K. Loke (Eds.), Rhetoric and Reality: Critical perspectives on educational technology. Proceedings ascilite Dunedin 2014 (pp. 629-633)

Cuban, L., (2001). *Oversold and underused: Computers in the classroom*. Cambridge, MA: Harvard University Press.

Cuban, L. 1986. *The classroom use of technology since 1920*. Teachers College Press, New York, NY.

Cutumisu, M., Blair, K. P., Chin, D. P., & Schwartz, D. L. (2015). Posterlet: A Game of Children's Choices to Seek Feedback and to Revise. *Journal of Learning Analytics*, 2(1), 49–71.

Design-Based Research Collective. (2003). *Design-based research: An emerging paradigm for educational inquiry.* Educational Researcher, 5-8.

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges.*International Journal of Technology Enhanced Learning*, *4*(5), 304-317.

Ferguson, R., & Buckingham Shum, S. (2011). Learning analytics to identify exploratory dialogue within synchronous text chat. In *Proceedings of the 1st International Conference on Learning Analytics & Knowledge,* (Banff, Canada)

Ferguson, R., Buckingham Shum, S., & Deakin Crick, R. (2011). EnquiryBlogger: using widgets to support awareness and reflection in a PLE Setting.

Ghislandi, P. M., & Raffaghelli, J. E. (2015). Forward-oriented designing for learning as a means to achieve educational quality. *British Journal of Educational Technology*, *46*(2), 280-299.

Gibson, J. J. (1977). The theory of affordances. In R. E. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing: Towards an ecological psychology*. (pp.67-82) Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Govaerts, S., Verbert, K., Duval, E., & Pardo, A. (2012). The student activity meter for awareness and self-reflection. In CHI'12 Extended Abstracts on Human Factors in Computing Systems (pp. 869-884). ACM. Hlupic, V., Pouloudi, A., and Rzevski, G. (2002). Towards an integrated approach to knowledge management: 'Hard', 'soft' and 'abstract' issues. *Knowledge and Process Management, 9*(2), 90-102.

Govaerts, S., Verbert, K., Klerkx, J., & Duval, E. (2010). Visualizing Activities for Self-reflection and Awareness. *In Proceedings of the 9th international conference on Web-based Learning*. (pp. 91–100). Springer, 2010.

Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Journal of Educational Technology & Society*, *15*(3), 42-57.

Haya, P. A., Daems, O., Malzahn, N., Castellanos, J., & Hoppe, H. U. (2015). Analysing content and patterns of interaction for improving the learning design of networked learning environments. *British Journal of Educational Technology*, *46*(2), 300-316.

Johnson, L., Becker, S., Estrada, V., & Freeman, A. (2014). *Horizon Report: 2014* Higher Education. Austin, TX: The New Media Consortium.

Janssen, J., & Bodemer, D. (2013). Coordinated computer-supported collaborative learning: Awareness and awareness tools. *Educational Psychologist*, *48*(1), 40-55.

Kolb, D.A. (1984). *Experiential education: Experience as the source of learning and development.* Englewood cliffs, NJ: Prentice Hall.

Kruse, A. N. N. A., & Pongsajapan, R. (2012). Student-centered learning analytics. *CNDLS Thought Papers*, 1-9.

Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning Learning Analytics with Learning Design. *American Behavioral Scientis*t, 57, 1439–1459. doi:10.1177/0002764213479367

Locke, E. A., & Latham, G. P. (2006). New directions in goal-setting theory. *Current Directions in Psychological Science*, 15(5), 265-268.

M'hammed Abdous, He, W., & Yen, C. J. (2012). Using data mining for predicting relationships between online question theme and final grade. *Journal of Educational Technology & Society*, *15*(3), 77-88.Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers & Education*,*54*(2), 588-599.

Macfadyen, L., & Dawson, S. (2012). Numbers Are Not Enough. Why e-Learning Analytics Failed to Inform an Institutional Strategic Plan. *Educational Technology & Society*, *15*(3), 149-163.

Manouselis, N., Drachsler, H., Verbert, K., & Duval, E. (2012). *Recommender systems for learning*. Springer Science & Business Media.

Mathan, S., & Koedinger, K. R. (2003). Recasting the feedback debate: Benefits of tutoring error detection and correction skills. In U. Hoppe, F. Verdejo & J. Kay (Eds.), *Artificial intelligence in education: Shaping the future of learning through intelligent technologies* (pp. 13-20). Amsterdam: IOS Press.

McAlpine, L., & Weston, C. (2000). Reflection: Issues related to improving professors' teaching and students' learning. Instructional Science, 28(5), 363-385.McNely, B. J., Gestwicki, P., Hill, J. H., Parli-Horne, P., & Johnson, E. (2012). Learning analytics for collaborative writing: a prototype and case study. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 222-225). ACM.

McKenney, S., & Mor, Y. (2015). Supporting teachers in data
journal of educational technology, 46(2), 265-279.Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher, 18*(2), 5-11.

McNely, B. J., Gestwicki, P., Hill, J. H., Parli-Horne, P., & Johnson, E. 2012. Learning analytics for collaborative writing: a prototype and case study. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (Vancouver, Canada, 2012). ACM, New York, NY, 222-225.

Melero, J., Hernández-Leo, D., Sun, J., Santos, P., & Blat, J. (2015). How was the activity? A visualization support for a case of location
Technology, 46(2), 317-329.

Miller, M. (2015). Leveraging CSCL technology for supporting and researching shared task perceptions in socially shared regulation of learning. (Unpublished doctoral dissertation) Univesity of Vicotria, Canada.

Norman, D. A. (2013). Design of Everyday Things: Revised and Expanded. New York: Basic Books. London: MIT Press (UK edition)

Nussbaumer, A., Hillemann, E. C., Gütl, C., & Albert, D. (2015). A Competence-based Service for Supporting Self-Regulated Learning in Virtual Environments. *Journal of Learning Analytics*, *2*(1), 101-133.

Oblinger, D.G. (2012). Let’s talk analytics. *EDUCAUSE Review,* July/August, 10-13

Postman, N. (1985). *Amusing ourselves to death*. New York: Penguin.

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of educational Psychology*, *95*(4), 667.

Rodríguez-Triana, M. J., Martínez-Monés, A., Asensio-Pérez, J. I., & Dimitriadis, Y. (2015). Scripting and monitoring meet each other: Aligning learning analytics and learning design to support teachers in orchestrating CSCL situations. *British Journal of Educational Technology*, *46*(2), 330-343.

Roll, I., & Winne, P. H. (2015). Understanding, evaluating, and supporting self- regulated learning using learning analytics. *Journal of Learning Analytics*, *2*(1), 7–12.

Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *3*(1), 12-27.

Segedy, J. R., Kinnebrew, J. S., Biswas, G. (2015). Using Coherence Analysis to Characterize Self- Regulated Learning Behaviours in Open- Ended Learning Environments. *Journal of Learning Analytics*, *2*(1), 13–48

Santos, J. L., Govaerts, S., Verbert, K., & Duval, E. (2012, April). Goal-oriented visualizations of activity tracking: a case study with engineering students. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 143-152). ACM.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.

Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional science*, *26*(1-2), 113-125.

Siemens, G. (2013). Learning Analytics: The Emergence of a Discipline. *American Behavioral Scientist*, *57*, 1380–1400. doi:10.1177/0002764213498851

Slade, S., & Prinsloo, P. (2013). Learning Analytics: Ethical Issues and Dilemmas. *American Behavioral Scientist*, *57*(March), 1–20. doi:10.1177/0002764213479366

Stolterman, E., & Wiberg, M. (2010). Concept-driven interaction design research. *Human–Computer Interaction*, 25(2), 95-118.

Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, 27(2), 237–246.

van Harmelen M. & Workman, D. (2012). Analytics for learning and teaching. CETIS Analytics Series 1(3).

van Leeuwen, A. (2015). Learning analytics to support teachers during synchronous CSCL: balancing between overview and overload. *Journal of Learning Analytics,* submitted for publication.

Vatrapu, R., Teplovs, C., Fujita, N., & Bull, S. (2011, February). Towards visual analytics for teachers' dynamic diagnostic pedagogical decision-making. In *Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (pp. 93-98). ACM.

Veenman, M. (2013). *Metacognition and learning: conceptual and methodological considerations revisited. What have we learned during the last decade?* Keynote given at EARLI 2013.

Verbert, K., Duval, E., Klerkx, J., Govaerts, S. & Santos, J.L. (2013). Learning analytics dashboard applications. *American Behavioral Scientist*, *57*(10), 1500-1509.

Verbert, K., Govaerts, S., Duval, E., Santos, J. L., Van Assche, F., Parra, G., & Klerkx, J. (2014). Learning dashboards: An overview and future research opportunities. *Personal and Ubiquitous Computing*, *18*, 1499–1514. doi:10.1007/s00779-013-0751-2

Visser, L., Plomp, T., Amirault, R. J., & Kuiper, W. (2002). Motivating students at a distance: The case of an international audience. *Educational Technology Research and Development*, *50*(2), 94-110.

Winne, P. H. (1995). Self-regulation is ubiquitous but its forms vary with knowledge. *Educational Psychologist*, *30*(4), 223-228.

Winne, P. H. (2011). A cognitive and metacognitive analysis of self-regulated learning. *Handbook of self-regulation of learning and performance*, 15-32.

Winne,P. H. (2000). Information processing models of self-regulated learning. In B. J. Zimmerman & D. H. Schunk (Eds.), Self-regulated learning and academic achievement: Theory, research, and practice. New York: Longman.

Winne, P. H., & Baker, R. S. J. d. (2013). The Potentials of Educational Data Mining for Researching Metacognition, Motivation and Self-Regulated Learning. *JEDM - Journal of Educational Data Mining*, *5*(1), 1–8. Retrieved from http://www.educationaldatamining.org/JEDM/index.php/JEDM/article/view/28

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), Metacognition in educational theory and practice (pp. 277– 304). Mahwah, N J: Erlbaum.

Winne, P.H., & Perry, N.E. (2000). Measuring self-regulated learning. In M. Boekaerts, P.R. Pintrich & Zeidner (eds), Handbook of Self-regulation. San Diego, CA: Academic Press.

Wise, A., & Vytasek, J. (2016). Learning analytics implementation design. In G. Siemens, Lang, C. (Eds.), Handbook of Learning Analytics. In review.

Wise, A., Zhao, Y., & Hausknecht, S. (2013). Learning analytics for online discussions: a pedagogical model for intervention with embedded and extracted analytics. *Conference on Learning Analytics,* 48–56. doi:10.1145/2460296.2460308

Wise, A., Zhao, Y., & Hausknecht, S. (2014). Learning analytics for online discussions: Embedded and extracted approaches. *Journal of Learning Analytics*, *1*(2), 48-71.

Wise, A. F. (2014). Designing pedagogical interventions to support student use of learning analytics. *Proceedins of the Fourth International Conference on Learning Analytics And Knowledge - LAK '14*, 203–211. doi:10.1145/2567574.2567588

Zeini, S., Göhnert T., Krempel L., & Hoppe H. U. (2012). The impact of measurement time on subgroup detection in online communities. In Proceedings of the 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, (pp. 389-394), Instanbul, Turkey: IEEE.

Zimmerman, B. J., & Schunk, D. H. (Eds.) (2001). *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Zimmerman, B. J., & Schunk, D. H. (Eds.). (2013). Self-regulated learning and academic achievement: Theoretical perspectives. Routledge.

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. American Educational Research Journal, 29(3), 663-676.

Using Learning Analytics to Identify Medical Student Misconceptions in an Online Virtual Patient Environment

Eric G. Poitras
University of Utah

Laura M. Naismith
University Health Network

Tenzin Doleck, Susanne P. Lajoie
McGill University

Abstract

This study aimed to identify misconceptions in medical student knowledge by mining user interactions in the MedU online learning environment. Data from 13000 attempts at a single virtual patient case were extracted from the MedU MySQL database. A subgroup discovery method was applied to identify patterns in learner-generated annotations and responses to multiple-choice items on the diagnosis and management of acute myocardial infarction (i.e., heart attack). First, the algorithm generated rules where single terms from the learner annotations were used to predict incorrect answers to the multiple-choice items. Second, the possible combinations of terms and their relevant synonyms were used to determine whether their inclusion led to better rates of prediction. The second step was found to significantly increase prediction precision and weighted relative accuracy, uncovering four misconceptions at a rate greater than 70%. These findings serve to inform the design of an adaptive system that tailors the delivery of formative feedback to promote better learning outcomes in the domain of clinical reasoning.

Introduction

Many students enter medical school with misconceptions that are resistant to change. A study of first- and second-year medical students showed that half of first-year students held one or more misconceptions prior to attending a course on the cardiovascular system, with this number only decreasing slightly in their second year of study (Ahopelto, Mikkalä-Erdmann, Olkinuora, & Kääpä, 2011). Students who held such misconceptions performed poorly on a related clinical reasoning task, supporting the fundamental role of knowledge in developing clinical reasoning expertise (Norman, 2005). In the absence of targeted instruction and detailed feedback, even more advanced medical students can struggle to debug their own misconceptions and restructure their knowledge appropriately (Boshuizen, van de Wiel, & Schmidt, 2012). Amongst practicing physicians, cognitive errors related to faulty prior knowledge or synthesis have been shown to contribute significantly to diagnostic error (Graber, Franklin, & Gordon, 2005). Furthermore, Graber et al. (2005) found that early errors influenced future errors, highlighting the need to identify and address misconceptions as early as possible in the learning process.

Misconceptions that lead to diagnostic errors can be challenging to examine in the clinical environment (Norman & Eva, 2010). In addition to providing valuable opportunities for students to practice and get feedback on their clinical reasoning skills, online learning environments can support instructors and researchers in tracking and logging user interactions to understand learning processes. Learning analytic techniques can be used to leverage the affordances of the large volumes of data generated by learners in these environments. The application of learning analytics to individualize instruction is referred to as learner modeling (Baker & Siemens, 2014; Ellaway, Pusic, Gallbraith, & Cameron, 2014; Ferguson, 2012; Kay, Reimann, Diebold, & Kummerfeld, 2013). Learner models allow instructional systems to capture and analyze user interactions in order to select and deliver the most suitable instructional content (Shute & Zapata-Rivera, 2012) and personalize the delivery of feedback (Feyzi Behnagh et al., 2014; Lajoie et al., 2013) to enhance learning outcomes.

In our previous research, we investigated the utility of a learning analytic technique known as subgroup discovery (Wrobel, 1997; Klösgen, 2002), a data-driven approach for generating rules that describe subsets of a population that are both sufficiently large and statistically unusual. We used this approach to classify learner interactions in another computer learning environment (Lajoie, 2009) where students learn to reason about virtual patient cases by formulating diagnoses and ordering laboratory tests. By examining the laboratory tests selected by learners in the context of diagnosing a virtual patient, we were able to generate rules that were suggestive of relationships between specific laboratory tests ordered and misconceptions in clinical reasoning (Poitras, Lajoie, Doleck, & Jarrell, 2016). However, the applicability of this approach for discovering knowledge from unstructured text-based data has yet to be determined. In the current study, we investigated the suitability of subgroup discovery methods for characterizing relationships between learner-generated text annotations and their subsequent responses to multiple-choice questions in the MedU online learning environment (Fall et al., 2005). Our aim was to examine the specific linguistic features of annotations associated with incorrect answers to identify common misconceptions, and to use such data to inform the provision of dynamic feedback in the context of problem-solving.

Detecting Misconceptions in Online Learning

Learning outcomes in online learning environments can be measured in terms of the degree of alignment between a learner model and an expert model of performance. Misconceptions and missing conceptions (Van Lehn, 1988) represent different types of discrepancies between these two models: missing conceptions refer to knowledge that the expert model contains that the student model does not, while misconceptions refer to knowledge that the student model contains and the expert model does not. Once detected, these discrepancies can be addressed in different ways. In the *model tracing approach* (Anderson, Boyle, Corbett, & Lewis, 1990), the learner model is continually compared against the expert model. If a discrepancy is identified, the learner is immediately and directly prompted to perform particular actions. In contrast, the use of *novice-expert overlay models* (Lajoie, Poitras, Doleck, & Jarrell, 2015) as a feedback mechanism puts the onus on learners to determine their next actions. The novice-expert overlay approach appears to be particularly effective in ill-structured domains where there may be multiple paths to obtaining the correct solution (Lajoie, 2003). Feedback that addresses learner misconceptions can be provided when erroneous solution paths are identified. Examples of the use of novice-expert overlay models to promote expertise in clinical reasoning include the Diagnostic Pathfinder (Danielson et al., 2007), the NUDOV system (Wahlgren, Edelbring, Fors, Hindbeck, & Stahle, 2006), and BioWorld (Lajoie et al., 2013, 2015). In this study, we extended the prior analytics methods we used with BioWorld to analyze learner actions in the more widely-used MedU online learning environment.

Virtual Patient Cases in MedU

MedU (www.med-u.org) has a current user base of 30,000 learners across North America and contains sets of virtual patient cases targeted at various specialties including pediatrics, internal medicine, family medicine, and radiology (Fall et al., 2005). A virtual patient case in MedU consists of a series of interactive HTML screens, or “cards”. Each card presents the learner with new information, including patient symptoms, current vital signs, electronic medical records, and consultation notes from other physicians. The sequential arrangement of the cards is intended to simulate how the condition of the patient evolves over time and in response to actions taken by the learner. The progression of patient symptoms is also made evident through explicit discussions of hospitalization, patient management plans, and treatment outcomes. The MedU environment embeds a number of tools to support learners in formulating their own differential diagnoses and treatment plans (see Figure 1). The *Navigation* sidebar allows learners to view the full set of cards in a case and navigate back to a previously viewed card. The *Tools/Resources* sidebar contains three free-text input fields where learners can record key findings, differential diagnoses and make annotations pertaining to the case.

MedU does not contain explicit learner or expert models. Instead, learner performance is analyzed through their responses to multiple-choice questions embedded throughout the case. These questions may address underlying knowledge and/or prompt learners to choose amongst a set of options for performing an action. The learner responses are highlighted as correct or incorrect and the appropriate actions that need to be taken are reviewed (e.g., “You should call a cardiology consult immediately”) and justified (e.g., “Urgency is critical as you want to prevent further myocardial damage…”). At the same time, the expert palette provides hints that are helpful in performing the correct action, as in a list of criteria to diagnose a left bundle branch block on an electrocardiogram (e.g., “The heart rhythm must be supraventricular in origin…”). Recent studies have sought to expand the nature of assessment in MedU. For example, Smith et al. (2016) developed a rubric for human tutors to assess students’ written case summary statements in MedU. We contend that learner annotations represent an important and as yet untapped source of data for detecting and addressing learner misconceptions.

Figure 1. A screenshot of the MedU user interface.

Research Objectives

The aim of this study was to detect learner misconceptions in the context of solving MedU virtual patient cases. Specifically, we applied a subgroup discovery method to search for common patterns in the learner-annotation data that were predictive of incorrect answers to multiple-choice questions. In doing so, our objective was to gain insights into the nature of misconceptions that mediate diagnostic performance and should therefore be targeted for corrective feedback. Our specific research questions included:

(1) Which linguistic features of learner annotations are associated with errors?
(2) Does the use of multiple terms increase prediction quality compared to single terms?
(3) Can the onset of misconceptions be detected?

Method

Materials

Dataset. We developed our analysis using the first case of the Structured Internal Medicine Patient Learning Experience (SIMPLE) set of MedU cases. In this case, Mr. Monson, a 49-year-old man, presents to the emergency department with an acute episode of chest pain, nausea and shortness of breath (Figure 1). The case walks the learner through the diagnosis and management of acute myocardial infarction (i.e., heart attack). This case consisted of 24 cards, 8 of which contained multiple-choice questions. With approval from the institutional review board at McGill University, we extracted anonymized student performance data for this case from the MedU MySQL database. The number of attempts for each multiple-choice question ranged from 12947 to 13262 (M = 13067, SD = 90). As the multiple-choice questions were designed such that multiple responses were required (i.e., "check all that apply"), we evaluated each item selection separately, for a total of 45 true/false items. For example, the first multiple-choice question asks students to select immediate treatment actions and provides four possible responses: (a) aspirin, (b) electrocardiogram (ECG), (c) sublingual nitroglycerin, and (d) a lab draw for cardiac troponins. These options were labelled as true/false items 1.1, 1.2, 1.3, and 1.4 for analysis purposes.

Sixty-seven percent of learners correctly identified acute myocardial infarction as a possible differential diagnosis (8786 correct, 4360 incorrect). This difficulty level suggested that some misconceptions were likely to be present.

Data scoring and coding. An example of a learner annotation made while using the Tools/Resources toolbar is shown below:

> *associated symptoms include nausea, mild epigastric tenderness dyspnea on exertion; relieved after 5 minutes of rest patient has history of tobacco use patient has family history of heart disease CXR normal no rubs, gallops, bruits RRR somatoform disorder pancreatitis GERD musculoskeletal - pulled muscle Angina not sx: no sweating, nausea, vomiting, radiating pain MI not in pain currently ischemic arterial disease - CAD pain in center of chest*

We extracted linguistic features of the annotations through a series of text pre-processing steps. The annotations were first isolated into terms by tokenizing the full string using non-characters such as spaces. For example, in the annotation above, the words "associated," "symptoms," "include," and "nausea" were each represented as separate terms. The terms were then transformed to lower case and filtered on the basis of length to exclude terms less than 2 characters. Any terms corresponding to stop-words such as *the*, *is*, *at*, *which*, and *on* were also excluded. A vector was created for each annotation using the complete set of terms extracted. Each cell in the vector contained a binary value representing term occurrence (i.e., true = 1) or non-occurrence (i.e., false = 0). The complete session-by-term vector was then pruned using a threshold of 10% in order to exclude terms with rare occurrences and reduce the dimensionality of the dataset.

Data analysis. Subgroup discovery involves an exhaustive search for relationships between learner behaviors, which consist of a set of predictor variables and a target variable of interest. In this study, the target variable in the dataset was the user selection for each of the 45 true/false items (i.e., binary value of 1 if selected or 0 if not selected). The subgroup discovery task was set to generate rules that account for incorrect responses to each item (i.e., true or false, depending on the item selected) on the basis of the linguistic features of learners' annotations to the Tools/Resources toolbar. For example, in the expert response to the first multiple-choice question, items a (1.1), b (1.2), and d (1.4) were selected. As our objective was to model incorrect responses, the target variables for the subgroup discovery algorithm were set to 0 (i.e., not selected) for items 1.1, 1.2, and 1.4, and 1 (i.e., selected) for item 1.3.

We applied the subgroup discovery algorithm to the MedU dataset using a two-step supervised approach. To facilitate analysis, the search task was constrained by a number of parameters. For example, the minimum coverage of a rule was set to 1% to ensure that errors in responding to the true/false items were prevalent enough to warrant intervention. In step one, the algorithm performed an exhaustive search across the dataset with the maximum depth of search set to 1, such that each rule contained a single-term *antecedent* from a learner annotation (e.g., nausea=true), that was associated with a *conclusion* (e.g., item 1.1=incorrect). We optimized for precision of rule detection. The number of task iterations was limited to select the set of 10 rules that was most precise in detecting errors for each true/false item. In step 2, the 10 antecedents from step 1 as well as relevant synonyms selected from an exhaustive word list generated from the text mining method were analyzed with the maximum depth of the search increased to 10. In doing so, the possible combination of terms and synonyms, up to a maximum of 10, were tested in order to determine whether a more precise solution could be obtained. All data analyses were performed in RapidMiner Studio, version 6.4 (https://rapidminer.com).

There are number of quality measures used in subgroup discovery (Herrera, Carmona, Gonzalez, & Jose del Jesus, 2011) to quantify the statistical novelty of a subgroup or rule (Konijn, Duivesteijn, Meeng & Knobbe, 2014). Duivesteijn and Arno (2011) note that selecting the right quality measure is often a difficult task; this choice is either driven by familiarity with the measures or based on default choice. For the present study, we calculated four common quality measures for each rule: Accuracy, Coverage, Precision, and Weighted Relative Accuracy (WRAcc; Lavrač, Flach, & Zupan, 2000). Accuracy was a measure of the proportion of instances where the rule made the correct prediction: (true positives + true negatives) / (all instances), while Precision represented the proportion of correct rule predictions amongst all cases where the antecedent term was found: (true positives) / (all positives). Coverage was calculated by dividing the number of annotations where the antecedent term was found by the size of the complete data set: (all positives) / (all instances). The WRAcc metric balanced both coverage and accuracy to assess the novelty of a subgroup relative to the entire population (Lavrač et al., 2000). WRAcc was calculated by subtracting the product of the probabilities that either the antecedent or conclusion were true from the probability that both the antecedent and conclusion were true: P (antecedent=true & conclusion=true) - [P(antecedent=true) X P(conclusion=true)]. WRAcc values can range from -.25 to .25, with values near 0 representing little novelty compared to the entire data set.

For each of the 45 true/false items, we calculated the average for each quality measure across the 10 rules generated for each step of the analysis. Paired *t*-tests were used to assess whether there were statistically significant improvements to any of the quality measures from step 1 to step 2 of the analysis. We calculated item difficulty by dividing the number of incorrect answers from the total number of item responses. The item difficulty value is thus independent from the quality measures used to appraise the rules generated through the subgroup discovery mining algorithm, and serves to corroborate the findings obtained from the analysis. In an attempt to facilitate early detection of misconceptions, we also mined the HTML content of the cards by extracting terms and matching them with a list of the antecedent terms generated by the subgroup discovery method.

Results

Summary Findings

After pruning to exclude rare terms, a total of 821 terms from the set of learner annotations were analyzed. Table 1 shows the descriptive statistics for each performance metric across all 45 true/false items. From step 1 to step 2 of the subgroup discovery algorithm, there was a significant increase in the precision of detecting errors, $t(44) = -10.73$, $p < .0005$ and in the weighted relative accuracy, $t(44) = -2.61$, $p = .01$, but no significant change to either accuracy, $t(44) = -1.62$, $p = .11$, or coverage, $t(44) = 1.42$, $p = .16$.

Table 1
Means and Standard Deviations for Subgroup Discovery Algorithm Quality Metrics (N = 45)

Metric	Step 1		Step 2	
	M	*SD*	*M*	*SD*
Accuracy	72.4%	19.3%	72.6%	19.4%
Precision	32.4%	21.7%	35.8%	22.5%
Coverage	1.9%	.70%	1.6%	1.2%
Weighted Relative Accuracy	.11%	.10%	.15%	.14%

Figure 2 and 3 shows the average precision and weighted relative accuracy values for each true/false item. Four items (2.3, 2.9, 3.4, and 3.5) were associated with average precision values greater than 70%. All of these items also had relatively high difficulty scores, ranging from 62% to 76%. In the following section, we elaborate on the most difficult item (i.e., 3.4) to illustrate the rule antecedents generated by the subgroup discovery task and the calculations of the different quality metrics.

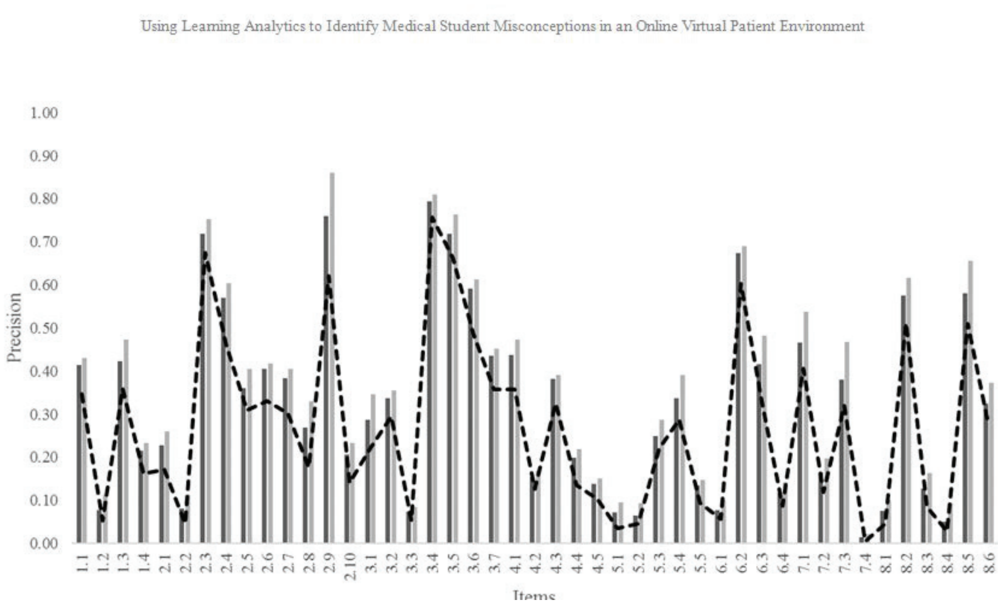


Figure 2. Average precision values for step 1 and step 2 across all 45 items for the Mr. Monson case.

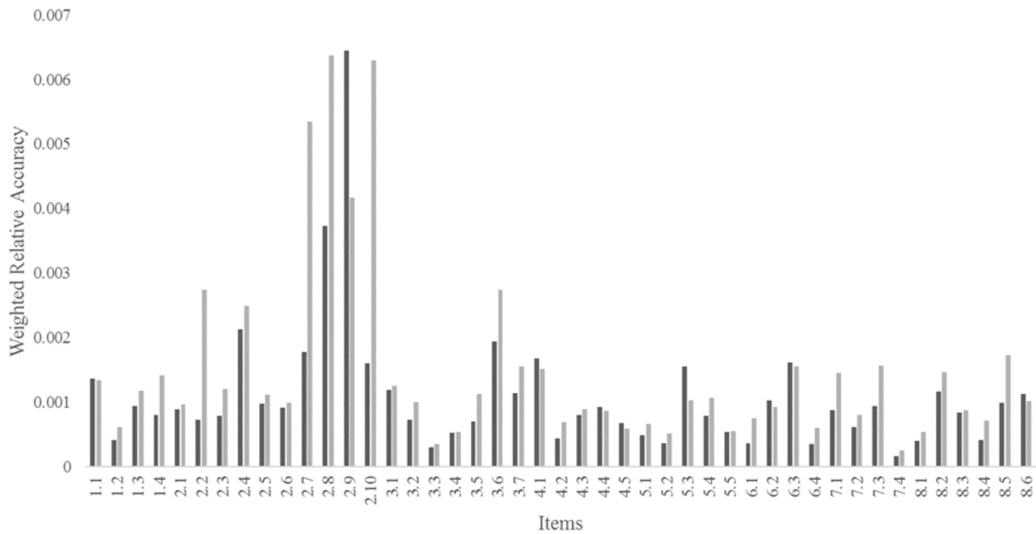


Figure 3. Weighted relative accuracy values for step 1 and step 2 across all 45 items for the Mr. Monson case.

Detailed Item Analysis

We detected a potential misconception related to the third multiple-choice question, which asked the learner to select the appropriate treatment for the management of a patient with ongoing chest pain due to unstable angina. Item 3.4 was especially difficult for learners. This item referred to the anticoagulant drug Heparin, which was selected (i.e. "true") in the expert response as being pertinent to the treatment plan. Of the 13073 learner responses to item 3.4 in the database, 9903 (76%) were classified as incorrect. The antecedents identified from the learner annotations in step 1 of the analysis included the terms: "segment," "neg," "reflux," "central," "felt," "carrying," "worse," "yrs.," "induced," and "dad." For example, of the 134 instances where the term "segment" was found in a learner's annotations, 108 responses to item 3.4 were also found to be incorrect, resulting in a detection precision of 80.6% (i.e., 108/134). As shown in Table 2, precision values increased in step 2 when multiple terms were used as antecedents.

In contrast to the high precision value, the accuracy of the term "segment" in detecting an incorrect response to item 3.4 was only 24.9% (i.e., [108 + 3144]/13073). This means that while mentioning "segment" was likely to be associated with an incorrect response to item 3.4, not mentioning "segment" was not necessarily indicative of a correct response to this item. As such, the mention of the antecedent term was a sufficient factor for inferring learner misconceptions, but not a necessary one. The coverage of this rule across the entire data set was 1.0% (i.e., 134/13073), meaning that misconceptions were typically rare amongst learners. Accordingly, the weighted relative accuracy was also low at .05% (i.e., 108/13073-[134/13073 X 9903/13073]).

A subset of the antecedents listed in Table 2 were found in the HTML content of the cards for this case. Item 3.4 appeared on the 12th card. The term "induced" was only mentioned on the first card (i.e., "Cocaine-induced chest pain") as a potential cause of chest pain related to the cardiovascular system. However, these terms figured prominently in the subsequent card where the patient complained about his symptoms:

*…So, while I was pulling carpet out and **carrying** it up the stairs, I noticed some discomfort in my chest. It was not really severe; it **felt** like some pressure. In retrospect, I think it might have been heartburn because I **felt** a little nauseated too…*

Table 2

Identified Rules from the Subgroup Discovery Task for Item 3.4

Antecedents	Tru	Fal	Cov	WRAcc	Pre
Step 1					
segment=true	108	26	0.01	4.97E-04	0.81
neg=true	149	37	0.01	6.20E-04	0.80
reflux=true	115	29	0.01	4.53E-04	0.80
central=true	110	28	0.01	4.18E-04	0.80
felt=true	156	41	0.02	5.18E-04	0.79
carrying=true	109	29	0.01	3.41E-04	0.79
worse=true	124	33	0.01	3.88E-04	0.79
yrs=true	138	37	0.01	4.16E-04	0.79
induced=true	141	38	0.01	4.13E-04	0.79
dad=true	421	114	0.04	0.001203	0.79
Step 2					
negative=false AND reflux=true AND yr=false	109	25	0.01	5.73E-04	0.81
neg=false AND negative=false AND reflux=true AND yr=false	108	25	0.01	5.55E-04	0.81
negative=false AND reflux=true	112	26	0.01	5.71E-04	0.81
neg=false AND negative=false AND reflux=true	111	26	0.01	5.52E-04	0.81
negative=false AND reflux=true AND segment=false AND yr=false	106	25	0.01	5.18E-04	0.81
negative=false AND reflux=true AND worse=false AND yr=false	106	25	0.01	5.18E-04	0.81
neg=false AND negative=false AND reflux=true AND worse=false AND yr=false	106	25	0.01	5.18E-04	0.81
negative=false AND reflux=true AND segment=false	109	26	0.01	5.15E-04	0.81
negative=false AND reflux=true AND worse=false	109	26	0.01	5.15E-04	0.81
neg=false AND negative=false AND reflux=true AND worse=false	109	26	0.01	5.15E-04	0.81

Notes: True Positives (Tru), False Positives (Fal), Coverage (Cov), Weighted Relative Accuracy (WRAcc), Precision (Pre)

Furthermore, these terms were also mentioned in the context of feedback provided to learners:

*…Most patients presenting with chest pain should have an ECG done immediately to look for ST **segment** abnormalities that indicate myocardial injury. ST **segment** elevations are present in a STEMI; in a NSTEMI, ST **segment** depressions may occur or the ST **segments** may be normal…*

Table 3 shows the terms that could be traced back to the HTML content of each card.

Table 3
Count of Item 3.4 Antecedent Term Mentions in the HTML Cards for the MedU Case

Antecedents	Case Card Number											
	1	2	3	4	5	6	7	8	9	10	11	12
segment=true		5									1	
neg=true												
reflux=true												
central=true												
felt=true		3										
carrying=true		1		1								
worse=true				1								
yrs=true												
induced=true	1											
dad=true		1										

Discussion

This study examined the application of learning analytic techniques to learner-generated annotation data in MedU, an online virtual patient environment. Based on data from a single case completed by over 13000 actual users, we were able to use the subgroup discovery technique to automatically generate rules to associate learner annotations with incorrect responses to 45 true/false items. From step 1 (single antecedent terms) to step 2 (multiple antecedent terms), we noted a significant increase in the precision and weighted relative accuracy of error detection, but no significant impact on either coverage or accuracy.

Responding to Ellaway et al.'s (2010) caution on the risk of high false positives when using large data sets, we purposely selected the parameters of our analysis to generate rules with a high precision of error detection (i.e., a low false positive rate). The high correlation between item difficulty and precision values obtained validated the choice of precision as the optimizing metric. We contend that the rules associated with precision values greater than 70% are likely to represent misconceptions, where the learner model contains knowledge that the expert model does not. For example, the results from step 1 suggest that an annotation of "reflux" was four times more likely to be associated with an incorrect answer to item 3.4 than a correct answer. All 10 rules generated in step 2 for this item also included the term "reflux" (see Table 2). Taken together, these results suggest that students may attribute Mr. Monson's symptoms to gastroesophageal reflux disease, a common condition that mimics myocardial infarction (Mayo Clinic, 2015). Conversely, the low overall accuracy of the rules can be attributed to a high false negative rate where the rule fails to detect incorrect answers. The low accuracy rate suggests that subgroup discovery may be less effective in detecting missing conceptions, where the expert model contains knowledge that the learner model does not. The low coverage and weighted relative accuracy values suggest that the detected misconceptions were relatively uncommon, and therefore unlikely to be detected by other means.

While automated techniques can successfully identify misconceptions, detecting their onset was considerably more difficult. By investigating the occurrence of the antecedent terms in the HTML content of the case, we were able to hypothesize about the nature of learner misconceptions. For example, the majority of the antecedent terms for item 3.4 appeared early in the HTML content of the case (see Table

3). Learners who used these terms in their annotations may have had difficulty updating their knowledge as the case evolved. Upon further investigation, we found that gastroesophageal reflux disease was a correct differential diagnosis for this case and was identified as such by 69% of learners (9088 correct, 4058 incorrect). Instead of a knowledge misconception, the inclusion of "reflux" as an annotation may reflect a tendency towards the cognitive error of premature closure, where other possibilities are not considered once an initial diagnosis is made (Graber et al., 2005). A complementary approach to searching the HTML content for the antecedent terms would be to investigate the context of the annotations in which the terms were mentioned. Both of these sources of information would allow domain experts to interpret the nature of the misconception in more depth.

Once identified, different approaches to providing feedback may be needed for different types of misconceptions. Model tracing approaches (e.g., Anderson et al., 1990) may be more effective for resolving misconceptions related to underlying knowledge, while novice-expert overlay approaches are likely to be more effective for supporting learners in self-regulating their cognitive processes while problem-solving (Lajoie et al., 2013).

Significance

Learning analytic techniques such as subgroup discovery provide an unprecedented opportunity to use data from real learners in authentic learning situations to better understand learning processes. This study illustrates how automated techniques can be used to detect learner misconceptions, formulate theory-based hypotheses about their sources, and inform the provision of personalized feedback to promote better learning outcomes.

Acknowledgements

This research was supported by a grant from Med-U. The content is solely the responsibility of the authors and does not necessarily represent the official view of Med-U.

References

Ahopelto, I., Mikkilä-Erdmann, M., Olkinuora, E., & Kääpä, P. (2011). A follow-up study of medical students' biomedical understanding and clinical reasoning concerning the cardiovascular system. *Advances in Health Sciences Education*, *16*, 655-668.

Anderson, J. R., Boyle, C. F., Corbett, A. T., & Lewis, M. W. (1990). Cognitive modeling and intelligent tutoring, *Artificial Intelligence, 42*, 7-49.

Baker, R., & Siemens, G. (2014). Educational data mining and learning analytics. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 253-272). Cambridge, UK: Cambridge University Press.

Boshuizen, H. P. A., van de Wiel, M. W. J., & Schmidt, H. G. (2012). What and how advanced medical students learn from reasoning through multiple cases. *Instructional Science*, *40*, 755-768.

Danielson, J. A., Mills, E. M., Vermeer, P. J., Preast, V. A., Young, K. M., Christopher, M. M., et al. (2007). Characteristics of a cognitive tool that helps students learn diagnostic problem solving. *Educational Technology, Research and Development, 55*(5), 499-520.

Duivesteijn, W., & Arno, A. (2011). Exploiting false discoveries–Statistical validation of patterns and quality measures in subgroup discovery. In *IEEE 11th International Conference on Data Mining (ICDM)* (pp. 151-160). Vancouver, BC: IEEE.

Ellaway, R. H., Pusic, M. V., Galbraith, R. M., & Cameron, T. (2014). Developing the role of big data and analytics in health professions education. *Medical Teacher, 36*(3), 216-222.

Fall, L. H., Berman, N. B., Smith, S., White, C. B., Woodhead, J. C., & Olson, A. L. (2005). Multi-institutional development and utilization of a computer-assisted learning program for the pediatrics clerkship: the CLIPP Project. *Academic Medicine, 80*(9), 847-855.

Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning, 4*(5-6), 304-317.

Feyzi Behnagh, R., Azevedo, R., Legowski, E., Reitmeyer, K., Tseytlin, E., & Crowley, R. (2014). Metacognitive scaffolds improve self-judgments of accuracy in a medical intelligent tutoring system. *Instructional Science, 42*(2), 159-181.

Graber, M. L., Franklin, N., & Gordon, R. (2005). Diagnostic error in internal medicine. *Archives of Internal Medicine, 165*, 1493-1499.

Herrera, F., Carmona, C. J., Gonzalez, P., & Jose del Jesus, M. (2011). An overview on subgroup discovery: Foundations and applications. *Knowledge and information systems, 29*(3), 495-525.

Kay, J., Reimann, P., Diebold, E., & Kummerfeld, B. (2013). MOOCs: So many learners, so much potential. *IEEE Intelligent Systems, 28*(3), 70-77.

Klösgen, W. (2002). Subgroup discovery. In W. Klösgen and J. Zytkow (Eds.), *Handbook of data mining and knowledge discovery*. New York: Oxford University Press.

Konijn, R., Duivesteijn, W., Meeng, M., & Knobbe, A. (2014). Cost-based quality measures in subgroup discovery. *Journal of Intelligent Information Systems*, 1-19.

Lajoie, S. (2003). Extending the scaffolding metaphor. *Instructional Science, 33*(5), 541-557.

Lajoie, S. (2009). Developing professional expertise with a cognitive apprenticeship model: Examples from Avionics and Medicine. In K. A. Ericsson (Ed.), *Development of professional expertise: Toward measurement of expert performance and design of optimal learning environments* (pp. 61-83). New York: Cambridge University Press.

Lajoie, S., Naismith, L., Poitras, E., Hong, Y., Panesso-Cruz, I., Ranelluci, J., & Wiseman, J. (2013). Technology rich tools to support self-regulated learning and performance in medicine. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies*. Amsterdam: The Netherlands: Springer.

Lajoie, S., Poitras, E., Doleck, T., & Jarrell, A. (2015). Modeling metacognitive activities in medical problem-solving with BioWorld. In Peña-Ayala (Ed.), *Metacognition: Fundaments, Applications, and Trends.* Springer Series: Intelligent Systems Reference Library.

Lavrač, N., Flach, P., & Zupan, B. (2000). Rule evaluation measures: A unifying view. In *Inductive Logic Programming, Lecture Notes in Computer Science, vol. 1634* (pp. 174-185). Berlin: Springer.

Mayo Clinic (2015). Diseases and Conditions: Heartburn.
http://www.mayoclinic.org/diseases-conditions/heartburn/in-depth/heartburn-gerd/art-20046483 (accessed July 17, 2015).

Norman, G. (2005). Research in clinical reasoning: Past history and current trends. *Medical Education, 39*(4), 418-427.

Norman, G. R., & Eva, K. W. (2010). Diagnostic error and clinical reasoning. *Medical Education, 44*(1), 94-100.

Poitras, E., Lajoie, S., Doleck, T., & Jarrell, A. (2016). Subgroup discovery with user interaction data: An empirically guided approach to improving intelligent tutoring systems. *Educational Technology & Society, 19*(2), 204-214.

Shute, V. J., & Zapata-Rivera, D. (2012). Adaptive educational systems. In P. Durlach (Ed.), *Adaptive technologies for training and education* (pp. 7-27). New York: Cambridge University Press.

Smith, S., Kogan, J. R., Berman, N. B., Dell, M. S., Brock, D. M., & Robins, L. S. (2016). The development and preliminary validation of a rubric to assess medical students' written summary statements in virtual patient cases. *Academic Medicine, 91*(1), 94-100.

Van Lehn, K. (1988). Toward a theory of impasse-driven learning. In H. Mandl & A. Lesgold (Eds.), Learning issues for intelligent tutoring systems (pp. 19-4 1). New York: Springer-Verlag.

Wahlgren, C.-F., Edelbring, S., Fors, U., Hindbeck, H., & Stahle, M. (2006). Evaluation of an interactive case simulation system in dermatology and venereology for medical students. *BMC Medical Education, 6*, 40.

Wrobel, S. (1997). An algorithm for multi-relational discovery of subgroups. In *Proceedings of the first European symposium on principles of data mining and knowledge discovery* (pp. 78-87). New York: Springer.

SECTION II: MOOCS, PSYCHOLOGICAL CONSTRUCTS, COMMUNICATION BEHAVIORS

Introduction to Section Two: MOOCs, Psychological Constructs, Communication Behaviors

Peter Shea, Editor-in-Chief, *Online Learning*
University at Albany, SUNY

This issue of *Online Learning* also contains four articles outside the theme of learning analytics. This section contains papers investigating MOOCs, a comparison of anxiety levels and the "imposter phenomenon" between online and classroom students, and a qualitative analysis of information behaviors among online students.

The first study is "Massive Open Online Courses (MOOCs): Participant Activity, Demographics, and Satisfaction" by Sara Shrader, Maryalice Wu, Dawn Owens, and Kathleen Santa Ana, of University of Illinois at Urbana-Champaign. In this paper the authors used clickstream and survey data to correlate demographic characteristics and levels of satisfaction with five separate clusters of participant activity that emerged from their analysis. They conclude that participants engage with MOOC courses in traditional and non-traditional ways, with some individuals choosing to participate in only one type of course activity, while others participate in multiple types of activity. Interestingly they found that non-completers are neither unsuccessful nor inactive. While the traditional patterns of participation (perhaps resulting in "course completion") may be preferred, non-traditional pathways are also valuable given that they also result high levels of satisfaction for those who engage this way.

In the next study "Global Times Call for Global Measures: Investigating Automated Essay Scoring in Linguistically-Diverse MOOCs" by Erin Reilly, Kyle Williams, Rose Stafford, Stephanie Corliss, Janet Walkow and Donna Kidwell of the University of Texas at Austin the authors also investigate Massive Open Online Courses. Specifically this paper investigates the whether the use of Automated Essay Scoring (AES) with non-native speakers of English is justified. Given the scale of MOOCs the use of such technologies is a potentially attractive, time-saving approach; however their reliability in the MOOC context is not well understood. The authors conclude that although non-native English speakers performed lower on free form essay and short answer English assignments overall, the AES system itself may differentially disadvantage non-native English speaking students resulting in lower scores than assigned by human graders. Clearly additional work is needed in this area given the vast number of international students currently involved in MOOCs and the need to accommodate even greater demand going forward.

Participation in graduate education can be a source of anxiety for many students. But do online environments present opportunities for reducing such anxiety and related negative psychological effects? These difficult questions are addressed in the next article, The Role of Social Influence in Anxiety and the Imposter Phenomenon by Christy Fraenza of Walden University. In this study the authors describe individual who suffer from the "imposter phenomenon" (IP) many of whom live their lives in constant fear of being discovered as less intelligent or competent than others believe them to be. The cycle of IP leaves the afflicted in a state of uncomfortable anxiety, especially in the face of imminent evaluations, such as those common to graduate education. One hypothesis guiding this study is that the source of anxiety triggering IP may be tied to social influence effects that are more salient in face-to-face settings. Does participation in online setting reduce the social influence that may result in anxiety tied to IP? Results indicated that graduate students in face-to-face settings had significantly higher IP scores than online graduate students. Results also indicated a significant, positive relationship between IP scores and anxiety scores. There are limitations to these findings of course, but the study raises very interesting issues for both classroom and online educators.

The final article in this section is Information Sharing, Community Development, and Deindividuation in the eLearning Domain by Nicole Cooke, of the University of Illinois at Urbana-Champaign. This study employs qualitative methods to provide a generative explanation of patterns of information interactions revealed in the written interactions of graduate students in an online learning community. Results provide a rich descriptive basis for understanding how students in an online Library and Information Science course create connections and build community over the span of a single term. The findings indicate that information behavior and sharing are not compartmentalized into strict cognitive and affective categories, but these domains complement one another to form a more comprehensive view of information seeking, utilization, and sharing, that contribute to the overall production of knowledge.

References

Cooke, Nicole A. (2016). Information sharing, community development, and deindividuation in the eLearning Domain. *Online Learning,* 20 (2), 244 - 260.

Fraenza, Christy B. (2016). The role of social influence in anxiety and the Imposter Phenomenon. *Online Learning,* 20 (2), 230 -243.

Reilly, Erin D.; Williams, Kyle M.; Stafford, Rose E.; Corliss, Stephanie B.; Walkow, Janet C.; & Donna K. Kidwell (2016). Global times call for global measures: Investigating automated essay scoring in linguistically-diverse MOOCs. *Online Learning,* 20 (2), 217 -229.

Shrader, Sara; Wu, Maryalice; Owens, Dawn & Santa Ana, Kathleen (2016). Massive open online courses (MOOCs): Participant activity, demographics, and satisfaction. *Online Learning,* 20 (2), 199 - 216.

Massive Open Online Courses (MOOCs): Participant Activity, Demographics, and Satisfaction

Sara Shrader, Maryalice Wu, Dawn Owens, Kathleen Santa Ana
University of Illinois @ Urbana-Champaign

Abstract

This paper examines activity patterns, participant demographics, and levels of satisfaction in multiple MOOC offerings at the University of Illinois at Urbana-Champaign from August 2012–December 2013. Using the following guiding questions: "Who are MOOC participants, how do they participate, and were they able to get what they wanted out of the course?" we have uncovered unique patterns of engagement that correlate with certain demographic characteristics. Our analysis employed both qualitative and quantitative methods, and serves as a model for further studies seeking to uncover the significance of participant activity within MOOCs.

Introduction

As enrollments soar and universities scramble to develop and deliver educational content to the masses, Massive Open Online Courses (MOOCs) have captured the attention and imagination of millions of people from around the globe. MOOCs have been designed so that anyone with an Internet connection can sign up for courses at little or no cost, thus eliminating traditional course-bound penalties for inactivity or incompletion. The University of Illinois at Urbana-Champaign (Illinois) was swift in joining the MOOC movement, partnering with Coursera in the summer of 2012. Initially, Illinois offered six MOOCs in topics that ranged from an introductory level course in *Sustainability* to a highly technical graduate level course in computer science known as *VLSI CAD* (Very-Large-Scale Integration Computer-Aided Design). This whirlwind of activity has prompted various stakeholders in the field of higher education to confront many overarching questions about the nature of MOOCs in order to better understand the significance of this emergent educational phenomenon.

Thus far a number of researchers have pondered key questions concerning who MOOC participants are and how they engage within MOOC platforms (Breslow, Pritchard, DeBoer, Stump, Ho,

& Seaton, 2013; Christensen, Steinmetz, Alcorn, Bennett, Woods, & Emanuel, 2013; Jordan, 2014; Wulf, Blohm, Brenner, & Leimeister, 2014). Our aim is to expand upon this endeavor in two ways. First, by presenting our approach to analyzing data associated with the first six MOOCs offered at Illinois in order to contribute to the growing scholarly discussion of how better to understand MOOCs and those who take them; second, to highlight the various patterns of activity displayed by our MOOC participants. Our focus on activity patterns is not to predict levels of course completion or incompletion, but instead to learn more about participant satisfaction based on differing activity patterns. Using clickstream and survey data, we correlate demographic characteristics and levels of satisfaction to five separate clusters of participant activity that emerged from our analysis.

Literature Review

The MOOC phenomenon calls forth an opportunity and need to examine new modalities of education, as well as to utilize different methods for analyzing the enormous amounts of data associated with MOOCs. Given the vastly different parameters that situate MOOCs from traditional modes of education—a context that is indeed both massive and open—researchers are confronted with the daunting task of conceptualizing ways to study MOOCs that allows for research replication, but maintain a readiness to discover promising approaches to gathering and analyzing data. Complicating the study of MOOCs even further is the differentiated nature of MOOCs where factors such as course design and discipline-specific attributes necessitate distinctive course goals and learning objectives, thus making cross-comparisons of data difficult.

Moreover, the ease of registering for MOOCs negates traditional course-bound rules such as formal drops and withdrawals, allowing MOOC participants to come and go as they please, to complete assignments or not. At the time of this study, MOOCs were mostly free of cost, offering a measure of autonomy for participants distinct from that of traditional degree-seeking students, as they are truly free to learn as they see fit because there are minimal financial and educational repercussions. Since involvement in MOOCs is solely at the discretion of participants and, for the most part, without penalty, the critical evaluation of traditional educational metrics is necessary in order to better contextualize the study of MOOCs (DeBoer, Ho, Stump, & Breslow, 2014).

As such, early rounds of MOOC research have called into question the efficacy of MOOC data, based on traditional metrics such as total enrollment numbers in determining completion rates (Chafkin 2013; Marcus 2013; Parr 2013; Pretz 2014). While MOOCs often enjoy a high level of initial enrollment, research shows that on average fewer than 10% of participants earn course completion certificates (Breslow et al., 2013; Ho, Reich, Nesterko, Seaton, Mullaney, Waldo, & Chuang, 2014; Koller, Ng, Chuong, & Chen, 2013). Concerns about low completion have been met with considerable critique (Carey, 2013; DeBoer et al., 2013; Ho et al., 2013; Vu & Fadde, 2014; Kizilcec, Piech, & Schneider, 2013), confirming the insufficiency of using total enrollment figures in determining completion rates. In turn, some researchers have examined intent for taking courses, which according to Koller et al. (2013) varies greatly, thus drawing some important distinctions about who registers for MOOCs and why.

While some authors have highlighted the potential of MOOCs as an effective means of democratizing higher education by increasing the access of non-traditional participants (Lewin, 2012; Wulf et al., 2014), demographic research has shown that MOOC users comprise a fairly homogeneous population. Typically, MOOC participants are young, well-educated males, living in developed countries, and have obtained higher levels of formal education (Christensen et al., 2013). Efforts have been made by institutions to target MOOCs toward certain populations (Mangelsdorf, Russell, Jorn, & Morrill, 2015), while MOOC service providers like Coursera have developed on-demand platforms to accommodate more flexibility in course offerings (Larson, 2014).

Other available MOOC research thus far has focused on identifying patterns of participant behaviors for both descriptive and predictive purposes. In particular, Kizilcec et al. (2013) used a longitudinal approach to track individual MOOC participants in three computer science courses. The authors identified four types of engagement, including completing, auditing, disengaging, and sampling. Similarly, Breslow et al. (2013) utilized predictive modeling in order to determine which course activities led to greater course success as defined by total number of points earned and persistence within a course.

Given the state of MOOC research, our paper seeks to examine the following questions:

1. How do our MOOC participants engage with the course material?
2. What are the demographic profiles of our MOOC participants?
3. Were participants able to get what they wanted out of the course?

Methods

In this paper we engaged in a retrospective analysis of the MOOCs offered at Illinois from August 2012–December 2013. Table 1 shows the differences in course length, course design, and course content.

Table 1: Summary of Course Length and Activities

Course	Duration	Activities	Enrollment
Introduction to Sustainability	8 weeks (3 offerings)	16 quizzes 46 video lectures 8 forum activities 1 final exam	91,325
Microeconomics Principles	8 weeks (4 offerings) 4 weeks (1 offering) 16 weeks (1 offering)	8 quizzes 89 video lectures 8 forum activities 8 project milestones	104,887
Introductory Organic Chemistry – Part I	8 weeks	8 quizzes in Coursera 8 quizzes in an external tool 100 video lectures	30,854
Intermediate Organic Chemistry – Part I	8 weeks	8 quizzes in Coursera 8 quizzes in an external tool 130 video lectures	14,434
VLSI CAD: Logic to Layout	10 weeks	8 quizzes 65 video lectures 4 programming activities 1 final exam	21,854
Heterogeneous Parallel Programming	8 weeks	7 quizzes 46 video lectures 7 programming activities	36,908

Over the span of two years, we gathered data both from within and outside of the Coursera platform. External to the Coursera platform is survey data that we gathered by administering two surveys to our course participants. At the beginning of each course, we sent out a questionnaire to all enrollees with questions about demographics, reasons for enrolling in the course, and what participants hoped to get out of the course. At the end of the course, a second questionnaire was sent, asking participants about their experience and satisfaction with the course (we did not collect survey data for the first offering of

Introduction to Sustainability or *Heterogeneous Parallel Programming*). Coursera also collected demographic data on course participants, and we merged that data into our own dataset to fill any necessary gaps.

Along with survey data, we collected the clickstream and event data recorded by the Coursera platform. By far the most granular, the clickstream data includes every click, of every participant, within a course site. Examples of clickstream data include forum and wiki views, as well as video views. The clickstream data allows us to know specifically which videos the participants watched, how many times they watched videos, and whether the videos were downloaded or streamed. If the videos were streamed, we can assess when participants paused, stopped, and/or restarted the videos. Clickstream data also contains information about the time and date of each click, the type of device and browser that was used to access various pages in the course site, and provides the IP address of every participant. The second source of data, the event data, contains information about video lectures viewed, quizzes taken, assignments submitted, dates of enrollment, the content of forum posts, quiz scores, and certificates earned.

As mentioned earlier, MOOC enrollment numbers cannot be meaningfully used as a central point of analysis in courses where participants are not held financially or academically accountable. The need to eschew total enrollment numbers in favor of using subtler points of analysis is continually reinforced in our approach to understanding the data. Take for example the following data (Table 2) from our first offering of Microeconomic Principles:

Table 2: Microeconomics Principles Enrollment Summary

Course	Total enrollment	Enrolled by the last day of the course	Enrolled after the last day of the course	Enrolled but never active	Enrolled and active during course period	Received a Certificate of Completion
Microeconomics Principles	50,676	50,375	301	28,107	22,569	2,233

By using “total enrollments” (50,676) as the denominator for determining the percentage of participants who completed the course by earning a “certificate of completion” (2,233), the course completion rate would be a mere 4.4%. However, if we chose to use participants who were “active during the course period” (22,569) as our denominator, the completion rate doubles to a completion rate of 9.9%.

We can fine-tune our analysis even further by making some decisions about what counts as course “activity.” If we defined activity broadly—clicking one time in the course site during the official start and end dates of a course—our active participant number is 22,569. However, by adjusting our definition of activity to include more nuanced time spent within a course, our results change yet again. Table 3 differentiates between activity for single day, for more than one day, and activity based on sustained, weekly clicks within a course:

Table 3: Microeconomics Principles Enrollment by Different Levels of Activity

Course	Active during the course period	Active for only a single day	Active for more than one day	Active every week of the course
Microeconomics Principles	22,569	8,019	14,550	2,261

By simply using “active for more than one day” as our denominator, completion rates jump to 15.3%. Conversely, by using “active every week of the course” as our denominator, completion rates balloon to 98.8%. Consequently, in order to paint a more nuanced picture of MOOCs, we opted to employ the more conservative “active for more than one day” denominator as the foci of our analysis in hopes of encapsulating a broad array of MOOC activity.

Likewise, we defined activity as a multifaceted action that includes various combinations of typical online course behaviors such as watching lectures, taking quizzes, submitting assignments, and participating in forums. Our understanding of activity is limited to the internal and external Coursera data outlined above, though presumably course participants may be engaged in relevant educational activity outside of the Coursera platform, including in-person study groups and social media gatherings. Specifically, we employed the following combinations of participant activity:

- only watching lectures
- only taking quizzes
- only engaging in forums
- watching lectures and taking quizzes, but not engaging in forums
- watching lectures and engaging in forums, but not taking quizzes
- taking quizzes and engaging in forums, but not watching lectures
- displaying activity in all three areas: watching lectures, taking quizzes, and engaging in forums

As such, our rationale for defining a MOOC participant in this manner—as someone who displays activity (as we have defined it) within a course for *more than one day*—is significant because final data reports will vary dramatically (as noted above) depending on the chosen denominator used to define participation.

Along with the need to better contextualize total enrollment data, researchers also need to clarify how they treat those who drop out of a course. In a traditional course structure, individuals who formally withdraw from a course tend to be classified as such, thus creating a separate category of analysis in regards to retention statistics. Yet, the concept of dropping out within the context of MOOCs makes little sense given the flexible signup process. Therefore, due to the conceptual problems involved with course withdrawals, our data includes individuals who discontinued their activity while remaining technically enrolled in the course, as well as those who intentionally un-enrolled themselves from a course.

Finally, some limitations must be acknowledged. Although we had 100% reporting on participant activity via web server logs, we had low response rates on the survey portions of our research where we ascertain participant demographic characteristics and course satisfaction. The survey response rates ranged from 3.1% (*Intermediate Organic Chemistry*) to 8.7% (*Introduction to Sustainability*) with an overall response rate of 5.3%. The response rates for people who are “Active for more than one day” ranged from 11.4% (*Introductory Organic Chemistry*) to 20.1% (*Introduction to Sustainability*) with an overall response rate of 14.0%. A low response rate is only a problem if it creates non-response bias (i.e. non-respondents are like respondents on variables of interest). To help ameliorate this bias, we calculated and employed non-response weights that use available information on participant activity levels regarding the true composition of the surveyed population in order to adjust the results to compensate for survey non-respondents. (Mandell, 1974) After adjusting for non-response bias according to activity patterns, we get a more accurate representation of the demographic distribution of MOOC participants.

After gathering and processing the clickstream, event, and survey data from each of our courses, we merged the data in order to uncover different patterns of behavior exhibited by MOOC participants by accounting for as much variation as possible.

Results

The following section illustrates the various activity patterns displayed by our course participants. Our analysis of the clickstream and event data reveals that participants engage with courses in traditional and non-traditional ways, with some individuals opting to participate in only one type of course activity, while others participate in multiple types of course activity.

Traditional and Non-Traditional Course Activity

Traditional course activity often consists of content being delivered through a lecture format, followed by some form of assessment that attempts to measure learning outcomes through an accumulation of points. Typically, in *traditional online courses*, videos are used to deliver content while computer and human-graded tests and quizzes serve as a central conduit of course assessment.

Progressing through a traditional online course, therefore, requires that participants watch a series of videos and complete various assessments over time. The process is structured in such a way that achieving pre-determined, instructor-created goals becomes the pathway for success. As a general convention for most educational schemas, completion or non-completion of the assessment pathway becomes the trigger that determines success. It is then *the quality of completion*, usually determined by a grading scale, which distinguishes high achievers from their less successful peers.

In contrast, non-traditional course activity takes a different form, where the concept of a structured pathway is muddled, and participants determine their own goals for success. Depending on an individual's pre-determined goals, success can mean watching only a handful of videos in order to learn a single concept, or success can mean watching some videos and taking multiple assessments. Since the openness of MOOCs allows for individuals to engage how they see fit, non-traditional course engagement necessarily becomes a valid path for success. Measuring success, however, becomes difficult since traditional course metrics such as completion do not hold sway outside of a pre-determined assessment pathway. Moreover, understanding success as a quality metric becomes difficult because there is no grading scale or other measure by which to determine achievement.

Participation patterns. Using "active for more than one day" as our denominator, activity within a course varies by course discipline, as does the kind of activity displayed by participants. However, there are also some consistent trends. Overall, we found two patterns of participant behavior, which include "single-activity" participants and "multi-activity" participants. Single-activity denotes participation in only one of the activities found within a course, while multi-activity encompasses participation in more than one course activity. Figure 1 and Table 4 illustrate the diverse activity types displayed by our Coursera participants.

Combining all courses, a significant percentage of participants only watched video lectures, accounting for nearly 29% (28.9%) of activity in all courses, though there are some clear outliers. Half of all participants in *Intermediate Organic Chemistry* only watched video lectures (49.4%), while 20.6% of participants in *Introduction to Sustainability* only watched videos with no other activity. Conversely, those who only submitted quizzes account for a fraction of participants within all courses (1.4%), with the highest percentage found in *Intermediate Organic Chemistry* at a mere 3.0%. Not surprisingly, forum only participants were also quite small (2.6%), with *Introductory Organic Chemistry* having the highest rates of forum participation (5.1%), and *Microeconomic Principles* having the lowest rates (1.6%).

Figure 1: Activity Type by Course

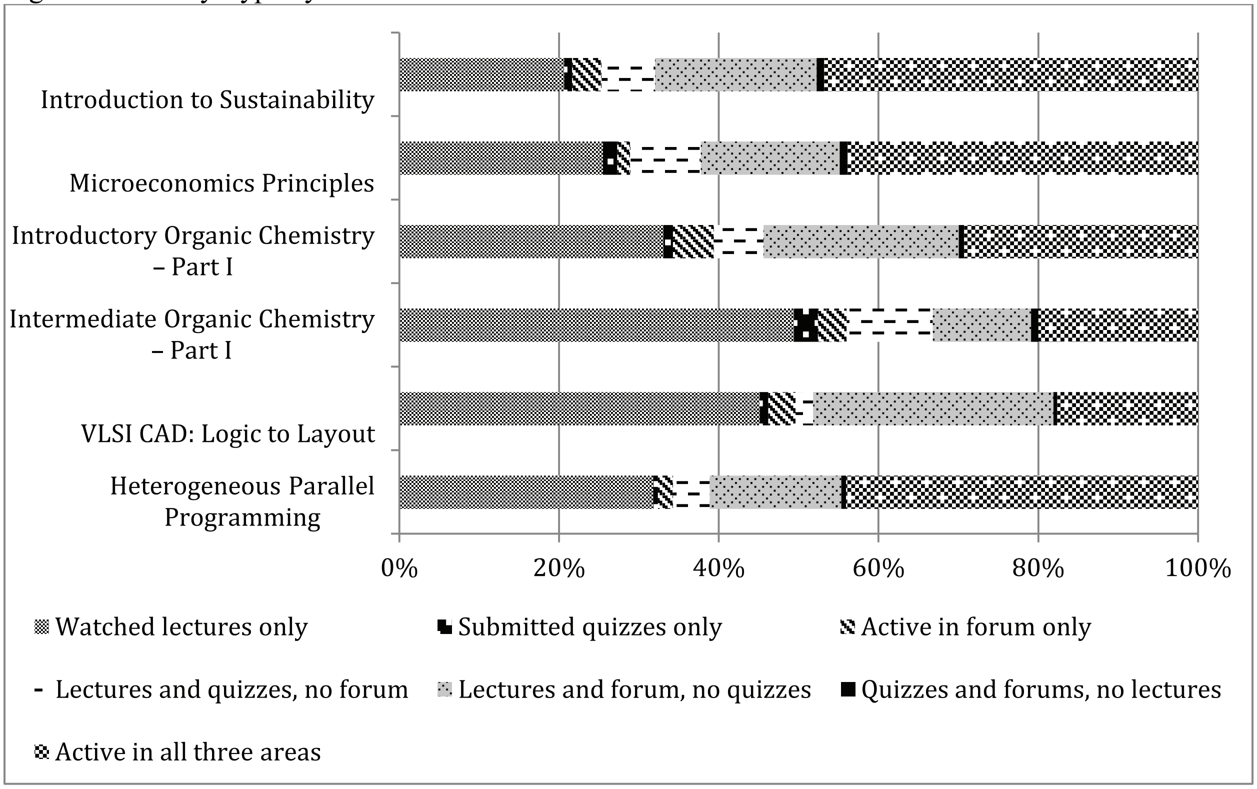


Table 4: Activity Type by Course

Course	Watched lectures only	Submitted quizzes only	Active in forum only	Lectures and quizzes, no forum	Lectures and forum, no quizzes	Quizzes and forums, no lectures	Active in all three areas
	%	%	%	%	%	%	%
Introduction to Sustainability	20.6	1.0	3.7	6.7	20.2	0.9	46.8
Microeconomics Principles	25.5	1.8	1.6	8.9	17.4	1.0	43.9
Introductory Organic Chemistry – Part I	33.0	1.2	5.1	6.2	24.4	0.7	29.2
Intermediate Organic Chemistry – Part I	49.4	3.0	3.6	10.8	12.3	0.9	20.0
VLSI CAD: Logic to Layout	45.1	1.1	3.4	2.2	30.1	0.5	17.6
Heterogeneous Parallel Programming	31.8	0.6	1.9	4.6	16.5	0.6	44.1
Overall	28.9	1.4	2.6	6.9	19.4	0.8	40.0

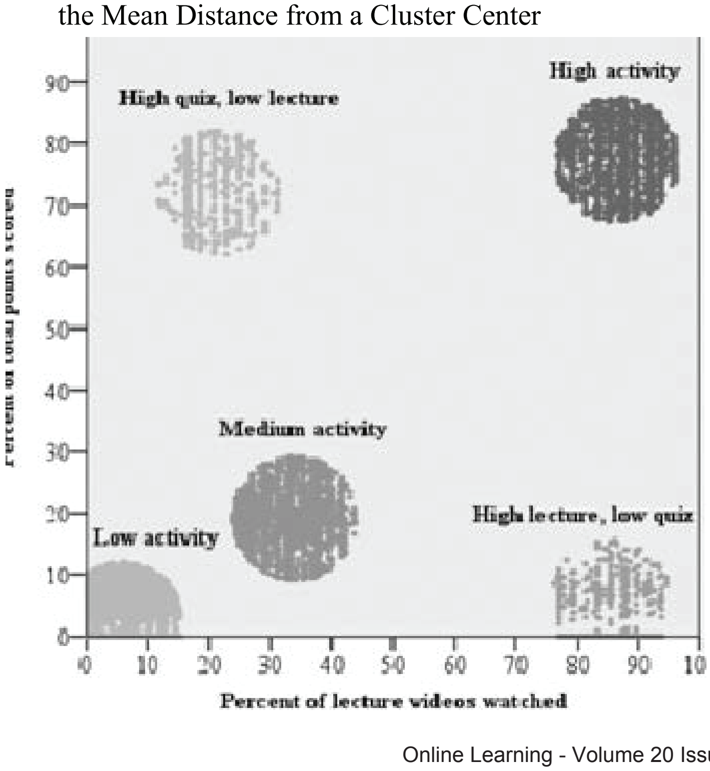
While more research is needed to understand why certain disciplines prompted stronger single-activity participation than others, it is clear that single-activity is a component of MOOC participation. This indicates that some participants did not use MOOCs as an all-encompassing, educative experience, but instead relied on discrete avenues for engaging with the course content. The openness of MOOCs necessarily invites this type of single-path participation, since participants are free to choose, in some respects, how they want to take a course. This type of participation closely mirrors the freedom of choice one engages in when using the Internet, where web users dictate how, when, and where they find and share information. In this way, single-activity participants used courses much like they would use the Internet by searching through a course in order to find a single activity of interest.

Engaging in multiple activities, however, accounts for higher percentages of participation for nearly all courses. On average for all courses, 40% of all participants engaged in all activity types (lectures, forums, quizzes). At 46.8%, participants in *Introduction to Sustainability* were the most multi-active, while *VLSI CAD: Logic to Layout* attracted the fewest multi-active participants (17.6%). "Dual-activity" participants, those who participated in a combination of two activity types, were most active by watching lectures and working in the forums and not taking quizzes (19.4%).

Due to the various combinations of single, dual, and multi-activity type, we applied *k*-means clustering analysis in order to better identify and understand patterns of participation. *K*-means clustering is an iterative technique used to group observations around the nearest means of two or more criterion variables. In our case we used "percent of videos watched" and "percent of total points scored" as criteria for clustering. Distance from the cluster means was calculated using simple Euclidean distance.

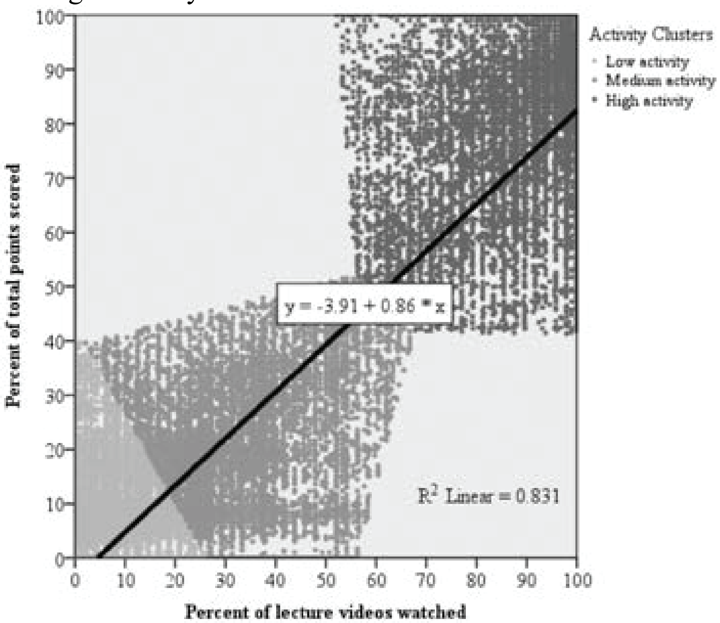
Using "active for more than one day" as our denominator, and modeling our two criterion variables of "percent of lecture videos watched" and "percent of total points scored" the following clusters highlight unique patterns of engagement among participants (forum participation was not included in this cluster analysis because it did not reveal any significant trends). These patterns suggest both traditional and non-traditional trajectories of course activity. It is important to note that our clustering model does not illustrate activity patterns on a time-scale; depicted in Figure 2 is the final outcome of participant activity at the end of a course.

Figure 2: Illinois Coursera Participants with the Mean Distance from a Cluster Center



In particular, our analysis shows that the "high" "medium" and "low" clusters emerged as candidates for proceeding along a traditional course trajectory, albeit at quite different levels of engagement within a course. Figure 3, also using *k*-means clustering, demonstrates a strong linear relationship (R^2=0.831) between those who watched videos and those who took quizzes, suggesting that these three groups (high, medium, and low activity) are actually one group on a single continuum.

Figure 3: Illinois Coursera Participants within the Low, Medium, and High Activity Clusters



For example, individuals in the "high activity" cluster demonstrated a strong level of engagement indicative of the traditional expectations that correspond to a traditional course setting. Participants who fell within this cluster watched a large percentage of videos (86% average), and earned a high percentage of quiz points (77% average). It is important to note, however, that engagement along a traditional path indicates little regarding the quality of such participation, and further research is needed to articulate the value of learning within a MOOC framework, relative to other traditional and non-traditional modes of learning. However, by traditional standards, participants in this cluster demonstrated success since 75% earned a certificate of completion.

In contrast, the "low activity" cluster includes individuals who engaged in a small number of course activities. The limited participation within this cluster, however, makes it difficult to extract any meaningful analysis without having access to more information. Perhaps the most intriguing group on the continuum is the "medium activity" cluster, since these individuals persisted longer than the "low activity" cluster, but fell short of the "high activity" cluster. Overall, 16% of participants in this cluster engaged in 25% of the activities within a course, while 62% completed half of the activities in a course. For the most part, those in the "medium activity" cluster have yet to differentiate what type of course participant they want to become. Although they are poised to travel down the "high activity" path, they are also in a position to go the non-traditional route of the "high quiz, low lecture" participants. Or, they could splinter off into the opposite direction, becoming part of the "high lecture, low quiz" group.

On the other hand, the "high quiz, low lecture" participants as well as the "high lecture, low quiz" participants appeared to be on a non-traditional course trajectory. For the "high lecture, low quiz" group watching videos is more important than completing quizzes, which suggests that these participants are not concerned with earning points and completing the course in the traditional sense. Likewise, the "high quiz, low lecture" group also engaged in the course the way they wanted to, using assessment as the primary means to interact with the content. Overall, individuals who chose non-traditional paths for course participation eschewed an educational experience based on achieving the pre-determined goals laid out for them. Instead, these participants sought an individualized path for themselves in terms of navigating the course.

It is important to note, however, that 54% of participants in the "high quiz, low lecture" group earned certificates of completion in the course. Our courses were primarily designed (with some exceptions) to reward participants who achieved a certain number of quiz points. This mirrors a common, and arguably traditional, educational practice of utilizing test scores as a primary means to award educational credentials. Ironically, in approaching the course in a non-traditional way, the activity of the “high quiz, low lecture” group reinforced an often critiqued practice of using tests as a primary means to incentivize learning.

Demographic patterns. In order to better understand the various profiles of participants who engaged with MOOCs through patterns of both traditional and non-traditional course activity, we ran individual logistic models for each of the five clusters with specific demographic characteristics pulled from our survey data and weighted by participant activity. Weights specific to each activity cluster within each course were constructed using the following formula:

$wt_n = (pp_n/ps_n)$ where:
wt_n = non-response weight for cluster n
pp_n = proportion of all participants in the course who were in activity cluster n
ps_n = proportion of all survey respondents who were in activity cluster n

After examining sex, age, education level, and employment status, the demographic composition of each cluster revealed some surprising insights, as well as some expected trends.

The only demographic characteristic that remained equal across all activity clusters is sex, where female and male participants are equally likely to be represented in each group. This finding is significant because some MOOC research has focused on the gender gap in certain MOOCs, illustrating the difference between female and male enrollments (Christensen, et. al, 2013). This descriptive enrollment gap can be seen in some of our courses as well. For example, although our overall course enrollments were split between males at 55% and females at 45%, *Introduction to Sustainability* and *VLSI CAD* did not follow this trend. The *VLSI CAD* population is highly skewed toward males (87.1%), whereas *Sustainability* is the only course that enrolled a majority of female participants (61.1%). Highlighting descriptive enrollment trends without additional analysis, however, paints a narrow and somewhat distorted view of what is really happening in MOOCs. Although females and males enrolled in different courses at different rates, their likelihood of course engagement at various levels, once enrolled in a course, is equal. So although a gender gap may exist upon enrollment, that gap closes once participants begin to engage in course activities.

Low activity. Predictably, the “low activity” cluster included the highest percentage of all age groups, education levels, and employment types. The high percentage of participants in this cluster was not surprising given other studies that have confirmed pervasive levels of low activity within MOOCs (Ho et al., 2014; MOOCs @ Edinburgh, 2013). However, participants who are 24 and younger are twice as likely to be in the “low activity” group compared to all other ages. This indicated that younger people tend to do less than their more mature counterparts.

Further research is needed to determine the factors that cause younger participants to disengage in a more pronounced manner, although some hypotheses can be made. For example, one can hypothesize that because younger participants are closer to the educational experience they are not looking to complete the course in the traditional sense. Students may be simply exploring different academic options in a quest to learn more about their own educational and professional interests. Whatever the reasons may be, once researchers have a more advanced understanding about this “low activity” cluster, decisions can be made on whether and how to adjust courses in order to foster sustained participant engagement for a larger number of registrants.

High activity. In contrast, participants in the “high activity” cluster revealed some stark differences in terms of age from the “low activity” cluster. For example, participants between the ages of 26-30 are 1.5 times more likely to be part of the “high” cluster than those between the ages of 18-24. Moreover, participants who are 30 and older are approximately twice as likely to be in this cluster than the 18-24 year olds. At 60 years old, however, participants are 4 times as likely to belong in the "high activity" cluster group than the younger 18-24 year olds. This indicates that older adults are more likely to persist (and even complete in the traditional sense) in courses than younger participants.

One possibility for such a wide age gap in the "high activity" cluster may be due to employment status. When looking at survey respondents who identified as either employed, unemployed, student, or retired, those who reported to be students were half as likely to be in the "high activity" cluster. The other three employment categories—employed, unemployed, and retired—are equally likely to be in this cluster. Students may be less likely to be in this cluster because they are busily working toward their degrees, leaving little time to focus on a MOOC. Or, the subject matter may be beyond a student's ability, leaving most students unable to keep up with the course material.

Moreover, participants with earned Masters or PhDs are approximately twice as likely to be in the "high activity" cluster, compared to those who have only completed high school or less. Presumably there are numerous reasons why individuals with advanced degrees are more likely to exhibit strong levels of engagement within a framework of traditional course activity. Perhaps their educational experiences have endowed them with the skills and cultural capital needed to persist in what is essentially a self-motivating environment. Or perhaps the sheer number of years that they attended school contributed to a trained mind regarding how one "takes a course." In any case, there is something to be said for the way in which established educational credentials provide a means for MOOC participants to maintain a high level of activity and achievement atypical of the less credentialed.

Medium activity. As discussed earlier, the "medium activity" cluster may be the most puzzling group to understand since individuals in this cluster did not display a clear path of activity. There were no statistical differences in this group in terms of sex, age, or education. The only interesting statistical difference is that unemployed participants are 1.4 times as likely to be in this cluster than their employed counterparts. Further research on the psychology of unemployment may help to explain this statistical anomaly. In terms of our analysis, however, this "medium activity" cluster continues to remain a mystery since it is unclear who these participants are, and why they displayed an inconspicuous level of course activity.

High lecture, low quiz. In terms of age, this group included the most striking statistical likelihoods. Overall, participants who are 50 years and older are four times more likely to be in this group than those ages 18-24. What's interesting to note is that starting at the age of 30, the likelihood of being in this group increases. This suggests that with increased age, individuals are more likely to prefer watching videos rather than taking quizzes or engaging with MOOCs through a traditional pattern. A similar pattern follows for educational level. Participants with a Masters and Bachelors are twice as likely to be in this cluster than those without a college degree. Moreover, participants who hold PhDs are over three times more likely to be in this cluster than those with lower educational credentials.

Finally, similar to the "high activity" cluster, the employment status of participants in the "high lecture, low quiz" group is composed mostly of non-students. In fact, students are half as likely to belong to this cluster, while retired participants are over twice as likely to be in this cluster. This reinforces a trend that seems to apply to older, well-educated, retired participants—not only is it highly likely that they will participate in courses through a non-traditional trajectory, but it is also highly likely that they will watch lectures without feeling a need to test their knowledge through predetermined assessments (such as quizzes).

High quiz, low lecture. Of all the clusters, the "high quiz, low lecture" group included the only equal distribution of all demographics—sex, age, education, and employment. Put another way, there is absolutely no difference regarding any of the four demographics in this cluster. This result was somewhat unexpected since it would be reasonable to hypothesize that this cluster would be composed heavily of students. It is quite conceivable to imagine that students would use MOOCs as a means to complement their current on-campus courses by taking advantage of free assessment activities. According to our analysis, however, this does not appear to be the case. Additional research is needed to determine what type of participants are more inclined to fall within the "high quiz, low lecture" cluster.

Participant satisfaction. As mentioned earlier, we administered two surveys at the beginning and end of each course. In the first survey, we asked two open-ended questions: *"What are your reasons for taking this course? What do you hope to get out of it?"* The function of asking these open-ended questions was twofold. First, we wanted to better understand the intent of Coursera participants, but also we hoped to discover more nuanced information about our participants that we could not predict fully by merely asking close-ended questions. While devising our course surveys, MOOCs were in their infancy (arguably, they still are) and little was known about how MOOCs would fit within the established system of higher education. As such, our research team was careful to construct survey instruments that allowed our research process to be informed by the survey data that we received, rather than imposing our preconceived notions about MOOCs onto the research process.

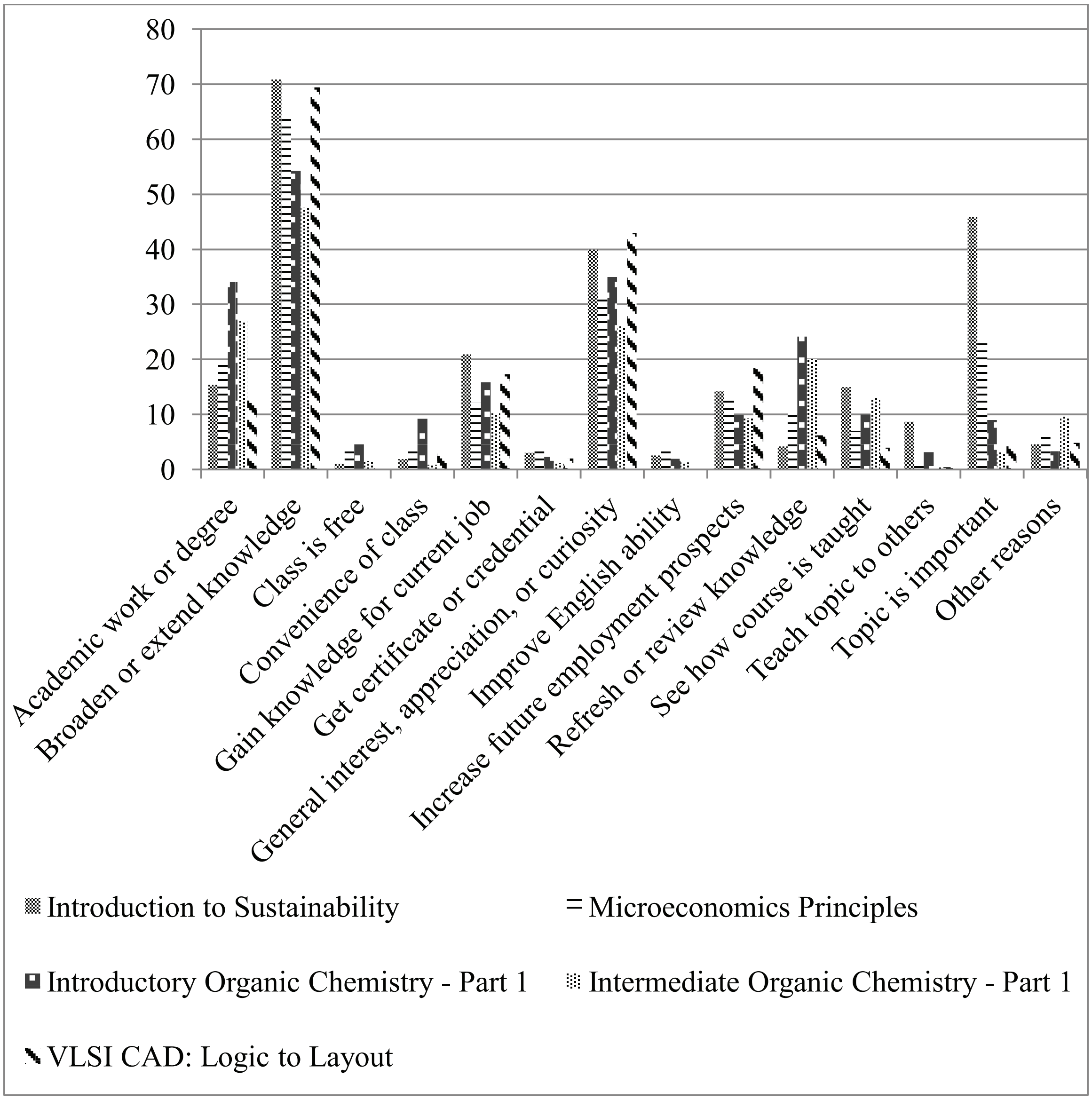
In order to analyze the open-ended survey questions, a coding process was established through thematic analysis resulting in the creation of a qualitative codebook. The codebook was then applied to 6,866 responses and coded by two people independently. All codes were then compared and analyzed for differences; all differences were reconciled between coders. This process resulted in 15,406 segments coded within the 6,866 responses. As a result, we uncovered some unique findings (Figure 4, Tables 5 & 6) with regard to the variety and novelty of responses offered by our course participants.

In particular, it was eye-opening to learn that some participants (albeit a small average, 2.8%) signed up for Coursera courses in order to improve their English ability. In designing our courses for this new, massive audience, we could not have imagined individuals would take courses for this reason. Further, it was surprising to learn that nearly 12% (11.7%) of participants signed-up for our MOOCs in order to see how the courses were taught. Presumably these participants were not interested in learning the content per se, but were simply curious about the buzz surrounding MOOCs. Another explanation is that they were interested in either designing their own MOOCs, or they hoped to get ideas for their own teaching.

Another important and surprising finding was that earning a certificate or credential was not an important reason for most individuals taking our Coursera courses (3.3%). Instead, the majority of participants offered reasons that are best described as falling under the umbrella of lifelong learning. In particular, a large percentage of survey respondents (35.6%) mentioned their general curiosity or interest in the topic, as well as their desire to broaden or extend their knowledge (65.6%).

In the second survey we asked the question: "*How much were you able to get what you wanted out of the course?*" Answer categories ranged on a five-point scale from: (1) *not at all* to (5) *to the fullest extent*. Through a comparison of means with an ANOVA test and a post hoc test of Games-Howell, three response patterns emerged. Not surprisingly, participants in the "low activity" and "medium activity" clusters reported the lowest levels of satisfaction at 3.04/5 and 3.08/5 respectively. These two groups were statistically different from the other three clusters, which reported a nearly one-point higher level of satisfaction. For example, participants in the "high activity" cluster reported significantly high levels of getting what they wanted out of the course (3.98/5). The "high quiz, low lecture" cluster was not far behind, reporting a level of satisfaction at 3.82/5. Participants in the "high lecture, low quiz" cluster indicated a satisfaction level of 3.44/5.

Figure 4: Reasons for Taking the Course



It is significant to note that all of the cluster groups reported above-average levels of getting what they wanted out of the course. Even the "low" and "medium" groups, which were somewhat of an enigma in terms of course activity and demographic composition, revealed a strong level of satisfaction with their course experience. The other three groups, which represented both traditional and non-traditional course trajectories, were highly satisfied with their course experience as well. Further, although none of the participants in the "high lecture, low quiz" cluster completed the course in the traditional sense (since receiving a certificate of completion in all of our MOOCs required that participants score a pre-determined number of points on course assessments), participants in this cluster felt that they were able to get what they wanted out of the course.

Table 5: Reasons for Taking the Course

Reason	Introduction to Sustainability		Microeconomics Principles		Introductory Organic Chemistry – Part I		Intermediate Organic Chemistry – Part I		VLSI CAD: Logic to Layout	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%
Academic work or degree	142	15.4	213	19.5	97	34.0	16	26.9	27	12.9
Broaden or extend knowledge	654	70.9	702	64.4	154	54.3	28	47.6	144	69.5
Class is free	9	1.0	38	3.5	13	4.6	1	1.6	1	0.7
Convenience of the class	18	1.9	37	3.4	26	9.2	1	0.9	7	3.2
Gain knowledge for current job	193	20.9	133	12.2	45	15.8	6	10.2	36	17.3
Get a certificate or credential	28	3.1	45	4.1	6	2.3	1	1.2	4	2.0
General interest, appreciation, or curiosity	373	40.4	334	30.6	101	35.7	16	26.2	90	43.3
Improve English ability	23	2.5	42	3.9	6	2.0	1	1.3	1	0.1
Increase employment prospects	131	14.2	138	12.7	28	10.0	6	9.4	38	18.5
Refresh or review knowledge	39	4.2	116	10.7	69	24.2	12	20.1	13	6.2
Teach the topic to others	80	8.7	11	1.0	9	3.2	1	0.4	1	0.4
See how course is taught	137	14.8	119	6.9	29	10.1	8	13.3	9	4.2
Topic is important	424	45.9	251	23.1	26	9.0	2	3.1	9	4.5
Other reasons	43	4.6	70	6.4	9	3.3	6	9.6	10	4.9

Table 6: Reasons for Taking the Course (Summary for All Courses)

Reason	Total	
	n	%
Academic work or degree	494	19.3
Broaden or extend knowledge	1682	65.6
Class is free	62	2.4
Convenience of the class	88	3.4
Gain knowledge for current job	413	16.1
Get a certificate or credential	84	3.3
General interest, appreciation, or curiosity	913	35.6
Improve English ability	72	2.8
Increase employment prospects	341	13.3
Refresh or review knowledge	248	9.7
See how course is taught	300	11.7
Teach the topic to others	102	4.0
Topic is important	712	27.8
Other reasons	138	5.4

Discussion and Further Considerations

The differentiated nature and duration of participant activity within MOOCs signifies a paradigm shift in how researchers are able to appropriate traditional metrics for understanding educational data. The porous structure of MOOCs complicates long-standing and universally accepted definitions associated most often with traditional course activity, minimizing the usefulness of benchmarks like completion or retention. In part, the varied demographic profiles of MOOC participants, as well as their intentions for taking courses, contribute to the enigmatic development of MOOCs as a new educational phenomenon.

For these reasons, our research approach has been to utilize different data sources (clickstream, event, surveys) in order to highlight the nuances of MOOCs to account for as much variation as possible. We identified various demographic characteristics of participants, and correlated those characteristics with the activity patterns displayed within our MOOCs. As a result, we uncovered activity patterns that reveal both traditional and non-traditional engagement within a course. The demographic characteristics of participants who engaged in the different course pathways varied, but overall satisfaction was consistently high regardless of one's chosen path.

The implications of our findings are important for a number of reasons. First, our results confirm the need to reconceptualize certain educational variables with regards to MOOCs. As discussed earlier, research has already indicated that using traditional educational metrics for understanding MOOCs is a fruitless endeavor (DeBoer et al., 2014). MOOCs are continually evolving as distinct organisms that do not necessarily fit within the archetype of traditional education. As such, the uncritical application of terms such as completion does little to further our understanding of MOOCs. According to our analysis, non-completers are not necessarily unsuccessful, nor are they necessarily non-active participants. While the traditional pathway may be preferred and even encouraged by a MOOC instructor, the non-traditional pathways elicit merit since these participants are reporting high levels of satisfaction.

In contrast, however, there is also a need to reaffirm some existing conceptions of traditional education with regard to MOOCs. Although MOOCs function within a space that is not fully aligned with traditional educational schemas, our findings show that nonetheless many participants choose to engage in MOOCs in a traditional manner. The “high activity” cluster clearly illustrates this phenomenon. Likewise, the “medium activity” cluster indicates that many MOOC participants instinctively attempt courses along a traditional trajectory. This instinct to “take a course” in a traditional way is, for many individuals, a natural pathway for engaging in sustained learning.

Part of the value in MOOC research, then, lies in further articulating notions of success in an online environment that continues, in some respects, to mirror components of a traditional online education, but has yet to fully develop a distinct identity of its own. MOOCs have rapidly evolved to meet the institutional needs of higher education as well as the populations served by these institutions. For many MOOC participants, earning a certificate of completion does not necessarily align with their intentions for taking a course, and it is important to further understand those specific MOOC populations who have redefined the course experience to fit their needs.

In rethinking MOOCs, new questions about the purpose and value of MOOCs as an educational model must be examined. For example, what utility should MOOCs serve for those single-activity individuals who are only concerned with watching videos or taking quizzes? If institutions view the value of MOOCs as updated versions of extension programs, or as an effective means to market the educational brand of an institution, then what should be done in order to meet these unique needs? On the other hand, if the goal is to provide pathways for those who have historically been marginalized from receiving a quality education, then what type of educational experience should be constructed? In asking (and attempting to answer) these types of questions, we need to be cognizant of whether or not these queries are mutually exclusive, and thus possibly contributing to a false narrative about MOOCs. In better understanding not only the MOOC phenomenon, but also continually examining the participant experience, researchers can effectively shape MOOC policy and practice as institutional goals and values continue to take shape.

Acknowledgements

We would like to thank Deanna Raineri, Rashid Robinson, Jane Blanken-Webb, Jake Astin, and Julian Martinez-Moreno for their contributions in the conceptualization of this paper. We’d also like to thank Jose Vazquez Cognet, Brian Ross, and Jason Mock who read earlier drafts, as well as the anonymous reviewers of *Online Learning* for their comments and recommendations.

References

Breslow, L., Pritchard, D., DeBoer, J., Stump, G., Ho, A., & Seaton, D., (2013). Studying Learning in the Worldwide Classroom: Research into edX's First MOOC. *Research & Practice in Assessment, 8*, 13-25.

Carey, K. (2013, December 12). Pay no attention to supposedly low MOOC completion rates [web log post]. Retrieved from http://www.edcentral.org/pay-attention-supposedly-low-mooc-completion-rates

Chafkin, M. (December 2013/January 2014). Udacity's Sebastian Thrun, Godfather of free online education changes course*. Fast Company, 181*. Retrieved from http://www.fastcompany.com

Christensen, G., Steinmetz, A., Alcorn, B., Bennett, A., Woods, D., & Emanuel, E. (2013). The MOOC phenomenon: Who takes massive open online courses and why? *Social Science Research Network,* Retrieved from http://dx.doi.org/10.2139/ssrn.2350964

DeBoer, J., Ho, A. D., Stump, G. S., & Breslow, L. (2014). Changing "Course": Reconceptualizing Educational Variables for Massive Open Online Courses. *Educational Researcher, 20*(10), 1-11.

Ho, A. D., Reich, B. J. F., Nesterko, S. O., Seaton, D. T., Mullaney, T. P., Waldo, J. H., & Chuang, I. (2014). HarvardX and MITx: The First Year of Open Online Courses, Fall 2012-Summer 2013. *Social Science Research Network,* Retrieved from http://dx.doi.org/10.2139/ssrn.2381263

Jordan, K. (2014). Initial Trends in Enrollment and Completion of Massive Open Online Courses. *The International Review of Research in Open and Distance Learning*, *15*(1), 133-160.

Kizilcec, R., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. *LAK '13: Proceedings of the Third International Conference on Learning Analytics and Knowledge, Leuven, Belgium,*170-179. doi 10.1145/2460296.2460330

Koller, D., Ng, A., Do, C., & Chen, Z. (2013, June). Retention and Intention in Massive Open Online Courses: In Depth. *Educause Review*. Retrieved from http://www.educause.edu

Larson, S. (2014, September). Online Education Gets Fast-Tracked with Coursera Classes On-Demand. *readwrite*. Retrieved from http://readwrite.com

Lewin, T. (2012, March 24). Instruction for masses knocks down campus walls. *The New York Times*. Retrieved from http://www.nytimes.com

Mandell, L. (1974). When to weight: Determining nonresponse bias in survey data. *The Public Opinion Quarterly, 38*(2), 247-252.

Mangelsdorf, S., Russell, J., Jorn, L., & Morrill, J. (2015, March). MOOCs for Wisconsin and the World. *Educause Review*. Retrieved from: http://www.educause.edu

Marcus, J. (2013, September). All hail MOOCs! Just don't ask if they actually work. *Time.* Retrieved from http://www.time.com

MOOCs @ Edinburgh. (2013). *MOOCs @ Edinburgh Report #1*. Edinburgh, Scotland: University of Edinburgh. Retrieved from http://hdl.handle.net/1842/6683

Parr, C. (2013, May 9). Mooc completion rates 'below 7%'. *The Times Higher Education*. Retrieved from http://www.timeshighereducation.co.uk

Pretz, K. (2014, 3 February). Low completion rates for MOOCs [web log post]. Retrieved from http://theinstitute.ieee.org/ieee-roundup/opinions/ieee-roundup/low-completion-rates-for-moocs

Vu, P., & Fadde, P. (2014). Rings of engagement: A model for MOOC enrollment. *Journal of Global Literacies, Technologies, and Emerging Pedagogies, 2*(3), 240-247.

Wulf, J., Blohm, I., Brenner, W. & Leimeister, J.M. (2014). Massive open online courses. *Business Information System and Engineering 6*(2), 111-114. doi: 10.1007/s12599-014-0313-9

Global Times Call for Global Measures: Investigating Automated Essay Scoring in Linguistically-Diverse MOOCs

Erin D. Reilly, Kyle M. Williams, Rose E. Stafford, Stephanie B. Corliss
Janet C. Walkow, Donna K. Kidwell
University of Texas at Austin

Abstract

This paper utilizes a case-study design to discuss global aspects of massive open online course (MOOC) assessment. Drawing from the literature on open-course models and linguistic gatekeeping in education, we position freeform assessment in MOOCs as both challenging and valuable, with an emphasis on current practices and student resources. We report on the findings from a linguistically-diverse pharmacy MOOC, taught by a native English speaker, which utilized an automated essay scoring (AES) assignment to engage students in the application of course content. Native English speakers performed better on the assignment overall, across both automated- and human-graders. Additionally, our results suggest that the use of an AES system may disadvantage non-native English speakers, with agreement between instructor and AES scoring being significantly lower for non-native English speakers. Survey responses also revealed that students often utilized online translators, though analyses showed that this did not detrimentally affect essay grades. Pedagogical and future assignment suggestions are then outlined, utilizing a multicultural-lens and acknowledging the possibility of certain assessments disadvantaging non-native English speakers within an English-based MOOC system.

Literature Review

MOOCS and the Promise of Open-Access Education

Massive open online courses, or MOOCs, have been controversial in the field of education, particularly higher education and educational research and assessment (Dolan, 2014; Watters, 2013). MOOCs are generally defined as large courses offered for free using open access materials and available to anyone with an internet connection. Although these courses do not typically offer credit from the providing institution, MOOC enrollees have the opportunity to earn completion certificates or "badges" (Liyanagunawardena, Adams, & Williams, 2013). Since 2011, millions of people around the globe have registered for hundreds of MOOCs delivered primarily through platforms such as edX, Coursera, and Udacity. An oft-reported goal for some MOOC providers is to allow access to educational materials for international learners who might not otherwise be able to take a course on a given subject due to distance, country, or socioeconomic status (Sandeen, 2013). Other online platforms have attempted using this model of "open access" to increase educational resources; however, there are still questions about the financial sustainability of such a movement (Yuan & Powell, 2013).

One of the major concerns of providing access to higher education in a global community is the need for academic literacy in English (Hamel, 2007). MOOCs—like many mediums for transmission of knowledge—are spreading institutional and educational content via the lingua franca of English through a global community of learners, with both negative and positive consequences (Kim, 2012). Given that the United States is still the primary purveyor and home for MOOCs, it is unsurprising that English is the predominant language for both the courses and technology developers for these platforms (Young, 2013). However, the current literature on distance education and online learning only occasionally considers potential difficulties for the international student learning experience due to linguistic issues.

Though MOOCs are typically globally accessible, they are not free of the cultural norms and potential Western biases that exist throughout global education. One of the most cited issues in global education is the historic (and expanding) role of English as a "gatekeeper" in education, employment, business opportunities, and promotion opportunities (Pennycook, 1999; Phillipson, 1992). Open-education researchers are beginning to call for a movement from "open access" to "open course" models that more thoughtfully incorporate local and global knowledge through student collaboration and participation in order to reduce the barriers for international students (Morgan & Carey, 2009). For instance, Liu, Liu, Lee, and Magjuka (2010), reporting on the impact of cultural differences for the learning experience of international students in an online environment, found that language is still one of the more predominate obstacles to ESL student participation online. The authors also note that a more culturally sensitive course design and instructor perspective may elicit higher levels of participation from international students, particularly for those who have difficulty with the English language.

Similarly, the goal of open access encounters, combined with the barrier of English language proficiency, suggests a need for greater understanding of global pedagogies. In an article detailing her experience as an international learner in online classes, Tan (2009) suggested that instructors integrate more cultural awareness and opportunities, such as the use of lecture videos, to enhance language proficiency. Additionally, Tan noted that detailed explanations of course expectations regarding grades and assignments in syllabi would be helpful in attracting international students who may be unsure of specific course expectations, a finding similarly noted by Liu and colleagues (2010). These factors lead to concerns regarding the consequences of this educational system; specifically, will MOOCs enhance and diversify student understandings of content, or privilege and homogenize the experiences of Western students and education?

Open-Ended Assessment and Non-Native English Speakers

For decades, the use of open-ended assessment for non-native English speakers has been considered both necessary and problematic within multiple academic disciplines (see Casannave & Hubbard, 1992). Open-ended responses and essays require learners to supply information as part of an assessment, as opposed to multiple-choice items that require learners to select a correct answer from a list of options. When thoughtfully constructed by an instructor, open-ended assessments evaluate higher-level learning and offer the opportunity for learners to receive detailed, personalized feedback (Attali & Burstein, 2006). Though these opportunities do not necessarily mean that students will use feedback, such assignments are meant to allow meaningful improvement in learning or performance (Furnborough & Truman, 2009).

In an effort to gain immediate feedback and assistance, students struggling with writing in a non-native language often utilize translation resources to improve their performance. For instance, educational researchers have suggested that the use of online translation systems is often utilized for immediate feedback when writing in a non-native language (Karabulut, Levelle, Li, & Suvorov, 2012). In fact, some researchers have suggested that increasing collaboration between MOOCs and universities from around the world has further pushed the boundary of online machine-translation technology to help non-native English speakers complete open-ended MOOC assignments (Clifford, Merschel, & Munné, 2013).

In terms of the impact of translational technologies, research has been inconsistent regarding whether this has a beneficial impact on course performance. For instance, Larson-Guenette (2013) found that learners' consistent use of online translation sites to check their work in another language was related to higher levels of active engagement with an assignment. However, the use of immediate open-ended feedback to enhance student motivation and self-regulation may be particularly important in the learning process, especially for distance language learners (e.g., Furnborough & Truman, 2009).

Automated Essay Scoring and Massive Open Online Classrooms

Due to the extremely large classes inherent in MOOCs, hand-scoring of open-ended assessments is often impracticable; consequently, some MOOC platforms have begun to integrate automated-essay scoring systems (AES) to grade open-ended assessments (Mayfield, 2013). Given that these AES tools are still in developmental phases, very little research has been conducted on the validity, perceptions, and best practices of AES systems embedded in MOOC platforms (Reilly, Stafford, Williams, & Corliss, 2014). Much remains to be learned about the specific scoring results of the tools. For instance, some critics of AES systems have argued that they are unable to accurately score higher-level writing tasks reflecting student performance expected at the college level (Condon, 2013; McCurry, 2010). Others note that AES systems do not accurately reflect scores that would have been given by an instructor, and may not process the nuances of writing in the way that a human grader can (Balfour, 2013). Additionally, the range of useful feedback for students engaging in AES assignments—a key part of the value of open-ended assessments—may also be a serious pedagogical issue.

There has also been a desire for a greater investigation of how AES learning activities can be utilized and structured in a way that does not disadvantage non-native English speakers (Chen & Cheng, 2008). One study performed through the Educational Testing Service found significant differences between the final human score and their e-rater system scores across language groups (Burstein & Chodorow, 1999). However, this difference did not significantly affect agreement between independent human raters, and the writing features examined by the e-rater were generalizable between native and non-native English speakers.

Another study by Guo (2009) on the Analytical Writing Assessment of the Graduate Management Admission Test (GMAT AWA) found that the AES tool provided by Intellimetric did not unfairly grade test-takers of different ethnicities, non-native English speakers, or students writing in a non-English language. Additionally, Dikli and Bleyle (2014) evaluated the use of the Criterion AES system within an ESL course reported that the AES system itself was useful to students attempting to refine their English writing skills and communication abilities. On the other hand, these AES systems are unique in that they have a long history of use and research on their validity and reliability, whereas newer AES systems such as those emerging in MOOCs have not been as thoroughly investigated or refined. Consequently, critics remain skeptical of the general ability of emerging MOOC AES systems to score student writing depending on different student language demographics and ability levels.

To contribute further to this line of inquiry, the following study utilized a single-course case study design to investigate the use of the edX AES system for a global, linguistically diverse audience. This research extends the work of authors such as Dikli and Bleyle (2014) by independently assessing an open-source AES system that allows for an instructor-created assignment and rubric. The edX AES system allows instructors to input a self-created rubric, which is then used by the instructor to grade 100 essays in order to train the machine-learning algorithm to assign rubric scores in the same manner. After the system is trained, the AES system assumes grading responsibilities for the remaining essays. For this study, the AES system gave back both rubric-category level and rubric-total scores. In the past, the edX AES rubric-system has shown adequate reliability and validity when compared to human graders, though the AES rubric-total tends to align best with shorter essays (Reilly, Stafford, Williams, & Corliss, 2014). Using information gathered from this course, including student grade data and survey responses, we sought to investigate the following questions.

Research Questions

1. Do self-reported levels of written and spoken English proficiency predict AES- and Instructor-essay scores?
2. Are non-native and native English speakers graded significantly differently on essays by the AES-grading system and human raters?
3. Are non-native English speakers graded significantly differently on non-essay assessments, when compared to native English speakers?
4. When controlling for English-language proficiency level, do students who use an online translator program receive higher scores than students who do not?

Method

Participants

Participants included MOOC student samples from an eight-week edX Pharmacy course, Take Your Medicine: The Impact of Drug Development. Each week consisted of two learning modules, which were lecture modules and videos with transcripts available in English. Each module was immediately followed by learning assessments that quizzed students on their module comprehension. In addition, the faculty developed a weekly learning lab for students where they applied what they learned that week by finding external online resources to answer the lab questions. In comparison to other assignments, labs particularly expanded beyond both the in-class content and traditional assessment tools by requiring students to research multiple-choice items through English-language pharmaceutical websites. This was an optional extra credit activity and the majority of the students participated.

Overall, 1,090 MOOC students completed the open-ended writing assignment. The mean age of students was 30.54 years and students were approximately equivalent regarding gender distribution (male = 48.5%, female = 51.5%). In terms of highest education level, 23% reported having a high-school degree, 32% a bachelor's degree, 24% a master's degree, 8% a doctoral degree, and 13% reported "other." The course was extremely diverse in terms of native language and English-proficiency level. In fact, a number of students in China set up satellite sites or small study groups, overseen by a course TA, in order to share and collaborate on translated class materials.

Of the students who completed the essay assignment, 35.26% identified English as their first language (EFL). Of the 64.74% English second-language (ESL) speakers, 85 distinct languages were reported, with Spanish (10.74%), Hindi (3.93%), and Portuguese (3.7%) as the top three non-English languages. Students also reported their proficiency (1 = not at all proficient to 5 = completely proficient) in both written English (M = 3.86) and spoken English (M = 3.87). Reported written and spoken proficiency scores were highly correlated (r = .80, p < .001). Post-course survey responses revealed that a goal for many students taking this MOOC was to gain proficiency in English (26%).

Procedure

The course essay assignment asked students to write a short-answer response of about 5 to 7 sentences reflecting on issues related to patient compliance with medical prescriptions. The students were then asked to answer five language-related questions regarding whether English was their first language (yes/no), what their first language was (write in), if they used an online translation program to write their essay (yes/no), and their level of proficiency with both written and spoken English. Additional course information was also collected, including course grades, post-lecture quiz grades, and lab assignments.

In total, 203 essays were randomly selected and de-identified for the purposes of this study. First, the original course instructor graded 100 essays within the edX platform to calibrate the AES system, which were not included in the study analyses. The instructor then graded an additional 203 essays according to the rubric used originally within the course. These additional essays were randomly selected and consisted of 65 EFL and 138 ESL speakers. The assignment utilized a rubric measuring four different areas: competence/understanding, support, organization, and content, with total scores ranging from 0 to 8 (see Table 1). When investigating differences across English proficiency levels, analyses were run using the total of rubric scores assigned by the AES and instructor.

Table 1. Essay assignment rubric categories and point allocation.

	Rubric Categories and Scoring			
	Rubric 1: Competence	**Rubric 2: Support**	**Rubric 3: Organization**	**Rubric 4: Content**
	Do you understand the potential health concerns and issues well?	**Do you provide evidence to support your arguments?**	**Is your essay clear and well organized?**	**Does your essay make appropriate reference to some of these keywords /phrases?**
Zero Points	Displays some serious misunderstandings of the issues it addresses; logic unclear, oversimplifies significantly, and gets some fundamental points wrong.	Little to no evidence provided to support idea.	Lack of organization makes underlying logic of argument difficult to follow.	No keywords or phrases used referring to relevant content.
One Point	Generally displays a good understanding of these healthcare issues, but on some finer points misunderstandings and confusions remain; gets the basics right, but oversimplifies a bit and misses some details.	Includes some supporting evidence but argument is not fully developed. Partially explores topic citing evidence from other sources.	Some organization makes underlying logic of argument possible to follow, but improvements can be made.	Makes appropriate reference to one or two keywords or phrases.
Two Points	Displays an excellent understanding of the issues it addresses, gets both the basics and details right. Shows an appreciation for logical structure and depth of issue.	Includes strong supporting arguments. Fully explores health-consequences using solid evidence.	Very organized and persuasive, explanations for argument are detailed and precise.	Makes appropriate reference to three or more keywords or phrases.

Results

The distributions of Instructor- and AES-scores were statistically and visually analyzed for normality. We found that the data substantially deviated from a normal distribution, as indicated by excessive levels of skewness (AES = -1.98, Instructor = -2.12) and kurtosis (AES = 3.78, Instructor = 4.59), as well as inspection of frequency distributions, boxplots, and Q-Q plots. Shapiro-Wilk tests also indicated that the score distributions significantly differed from normality (AES = 0.67, $p < 0.0001$; Instructor = 0.64, $p < 0.001$). This non-normality was likely due to the eight-point scale used in calculating total essay scores; therefore, we used only non-parametric analyses. Means and standard deviations for rubric level and total scores by grader and English-speaking category are provided in Table 2.

Multiple analyses were conducted in order to determine the nature of the relationship between the AES scoring system and the instructor's grading. Spearman correlations indicated that there were significant positive correlations between essay scores and reported English proficiency, such that greater levels of spoken English proficiency ($r_s(202) = .24$, $p < .001$; $r_s(202) = .28$, $p < .001$) and written English skills ($r_s(203) = .26$, $p < .001$; $r_s(203) = .27$, $p < .001$) predicted both higher AES-graded and instructor-graded essay scores, respectively. Data collected on students' self-reported English language proficiency (first language information, written ability, spoken ability, and use of an online-translator program) was utilized to investigate grading differences by AES and Instructor. Wilcoxon-Mann-Whitney tests were used as a non-parametric version of an independent samples t-test. Findings indicated that MOOC students reporting a non-English first language were scored significantly lower than EFL students by both the AES total (z = 2.94, $p < .01$) and Instructor total (z = 2.97, $p < .001$).

Table 2. Average scores assigned by the AES and Instructor for EFL and ESL students.

	AES Scores		**Instructor Scores**	
Variable	**EFL**	**ESL**	**EFL**	**ESL**
Rubric 1	1.86 (0.43)	1.70 (0.60)	1.92 (0.32)	1.84 (0.46)
Rubric 2	1.82 (0.46)	1.57 (0.68)	1.92 (0.27)	1.75 (0.53)
Rubric 3	1.82 (0.53)	1.67 (0.62)	1.89 (0.31)	1.71 (0.49)
Rubric 4	1.88 (0.38)	1.72 (0.56)	1.78 (0.45)	1.66 (0.56)
Total	7.37 (1.41)	6.67 (1.92)	7.52 (1.13)	6.96 (1.61)

*Note. EFL = English-First Language students (n = 65). ESL = English-Second Language students (n = 138).

Percent agreement between the Instructor and AES was calculated to describe the overlap in the scorers (AES and instructor) on each rubric point and the total score. These were calculated across all students and for EFL and ESL students separately (see Table 3). We also analyzed χ^2-statistics to determine whether the rate of Instructor-AES agreement differed by English-language category. Agreement on individual rubric scores ranges from 77% to 85% EFL students and from 72% to 78% for ESL students. Though percent agreement between Instructor and AES rubric-category scores was consistently higher for native English speakers than ESL speakers, these differences were not statistically significant. However, the percent agreement of total scores was significantly higher for EFL students than ESL students ($\chi^2 = 7.64$, $p < .01$), being separated by a margin of over 20% (EFL = 69%; ESL = 49%).

Table 3. Percent Agreement between Instructor- and AES-Scores

Score	Rubric 1	Rubric 2	Rubric 3	Rubric 4	Total
EFL	83.08%	84.62%	81.54%	76.92%	69.23%
ESL	77.54%	74.64%	72.46%	72.46%	48.55%
χ^2	0.83	2.55	1.96	0.46	7.64^{**}

Note. EFL = English-First Language students. ESL = English-Second Language students. χ^2 = test for association between English-language category and percent agreement, $^{**}p < .01$

As the percent agreement may have been higher for native English speakers due to the high abundance of perfect scores and lower variability in this group, we analyzed several other inter-rater agreement indices that account for the probability that the AES and instructor scores would agree due to the prevalence of scores assigned. The chance-corrected coefficients analyzed were Cohen's κ, Scott's π, and Krippendorf's α, which can all be interpreted as the percent agreement between the instructor and AES above that which is expected by chance. These coefficients produced consistent results and indicated that there was slightly greater instructor-AES agreement for EFL students than for ESL students, after accounting for the likelihood of ratings agreeing due to chance (see Table 4). However, the agreement rates of the English categories were very similar, and lower than would be considered acceptable in most social science applications, with the instructor and AES agreeing 21%-25% of the time after deducting the probability they would agree due to chance.

Table 4. Chance-corrected agreement between instructor and AES-Scores

	κ	π	α
Native English Speakers	0.24	0.24	0.25
Non-Native English Speakers	0.22	0.21	0.22

κ = Cohen's Kappa. π = Scott's Pi. α = Krippendorff's Alpha.

Wilcoxon signed rank tests (non-parametric repeated measures t-tests) were used to compare the average scores of students who did and did not use an online translator for essay writing, after matching students on English written proficiency and English-as-first-language status. Results suggested that, when controlling for English language proficiency, there was not a significant difference in AES total scores between students who used an online translation program versus those that did not (S = 309, p = .07). Wilcoxon-Mann-Whitney tests revealed that, for native and non-native English speakers, average post-lecture comprehension quiz scores (z = 1.12, p = .26) and average course grades (z = 1.45, p =.15) did not significantly differ, although lab grades (z =2.75, p <.01) were significantly higher for native English speakers.

Discussion

As MOOCs continue to serve a growing global audience, the need for linguistically sensitivity and globally applicable assessments of learning will also continue to grow. Our findings suggest that non-native English speakers are graded significantly lower by the AES grading system as well as by the instructor, when compared to native English speakers. Additionally, the results from this study show a positive and significant relationship between English proficiency and essay scores in total, with higher written and spoken English-language proficiency correlating with higher AES rubric scores and instructor scores.

This indicates that students who rate their English-language proficiency level more highly tend to receive better scores on their essays from the AES system as well as the instructor, while students who rate their proficiency level lower tend to receive lower essay scores from both graders. On the other hand, the level of score-agreement between instructor and AES grading was higher for native English speakers as compared to ESL students, even when reanalyzed to correct for chance agreement; this suggests that the AES system may in fact be less valid and comparable to human grading when scoring non-native English speakers.

Taken in conjunction, these findings suggest that although non-native English speakers performed lower on this assignment overall, the AES system itself may differentially disadvantage non-native English speaking students. Additionally, research suggests that open-ended assessment itself benefits native English speakers, as has been shown in previous literature on the use of essays in global distance learning (Goodfellow & Lea, 2005). Our finding that native English-speaking students performed significantly better on the lab exercises, which required online and English-based research on external websites, suggests that other types of assessment that draw on literacy-based skills may also disadvantage students with a non-English first language. Consequently, the necessity of multiple forms of assessment, as well as other low-stakes open-ended writing assignments (e.g., discussion boards), should also be assessed and utilized in order to support students in freeform learning experiences.

Finally, the findings from this study did not support the hypothesis that the use of an online translator would result in higher AES system scores after controlling for English-language proficiency level; in fact, there was no difference between students' rubric total scores in regards to their use of an online translation program. This suggests that the use of machine translation for non-native English speakers does not significantly impair, nor enhance essay quality, which supports research suggesting that web-based translators are not necessarily effective in translating text into another language in a way that aids in the quality of essay assignments (Williams, 2006). Further research in the area of AES systems and online translation programs may shed light on the strengths and shortcomings of using AES grading in non-native speaker populations as an assessment tool, given the increasing availability of free translation software.

Together, these results suggest that differences may exist between native and non-native English speakers when students are graded by AES systems, which is a clearly complex problem when examining the intentions of MOOC audiences. MOOCs have been hailed as an educational resource for learners outside of the United States to have quality access to educational resources and valuable learning experiences (Byerly, 2012; Meyer & Zhu, 2013). However, non-native English speaking populations appear to be at a disadvantage because of their language proficiency on certain assessments, and thus may not be well-served by MOOCs that intend to use essay or research-oriented assignments for high-stakes testing in their courses. MOOC issues must be examined in terms of equity and adequacy of global education, such that students might experience "linguistic gatekeeping" if such assessment types are used as part of high-stakes testing in global courses (Phillipson, 1992). As Cushing Weigle (2013) notes in her study of AES systems and language diversity, the benefits of implementing a more expedient system for scoring writing should be weighed against the potential for marginalizing non-native English speakers within an English-based education system.

Multicultural and Pedagogical Considerations

In terms of fairness in learning and assessment, these findings have some prospective practical applications and future research suggestions. When the group of learners MOOCs are attempting to reach are also placed at a disadvantage for the purposes of evaluation and grading, educators must re-examine the usefulness and applicability of such assessments. For example, some educational researchers have suggested a global-education paradigm shift from emphasizing "native-like fluency" in English, and instead promoting "reasonable competence" (Methitham, 2009). Further research on the AES tool in other disciplines may help MOOC instructors and instructional designers better understand the ways in which this tool could be used to support student learning, as well as allow for refinement of the AES system algorithm to more accurately assess non-native English speakers' writing ability (Cushing Weigle, 2013; Dikli & Bleyle, 2014).

Culturally-sensitive research methodologies suggest the importance of taking language proficiency into account throughout the assessment and course-design process, in order to address potentially problematic assessments and grading issues (Uzuner, 2009). After engaging with the tool, it has been suggested that instructors hand-grade student answers after using the AES system to better determine potential rubric issues and levels of subject mastery (Walkow & Reilly, 2014). Ideally, students would respond in their native language, though this is difficult to accommodate when dozens of countries and languages are represented. One possible option would be identifying non-English language options (Spanish, Hindi, Arabic, etc.) for future AES systems to provide a more inclusive environment.

On the other hand, many students reported gaining greater English proficiency as a goal for this MOOC. Thus, creating formative opportunities for students to engage in English-based open-ended assignments might also be pedagogically useful. Such assignments in globally-available open courses may allow students to more actively achieve their own learning goals through self-regulation, beyond specific course outcomes (Peters, 2002). In order to improve access to course content globally—as opposed to "educating the educated" (Hollands & Tirthali, 2014, p. 13)—it is important that faculty, course developers, and platform programmers take into account the diversity of learners in MOOCs, and identify ways to help them meet their course goals.

Limitations

Several limitations for this study were present. First, language information was only acquired for students who completed the essay assignment. As Dikli (2006) has pointed out, AES systems that utilize prior calibration for grading accuracy can be only as good as what they learn from the initial calibration sample. Consequently, it is likely that many non-native English students opted out of this assignment, and thus did not have their data included in this analysis. Second, plagiarism was not accounted for by either the instructor or the AES system in this scoring set, though the instructor did find multiple plagiarized essays that were scored highly by the AES system. Future MOOC AES systems should attempt to incorporate plagiarism software in order to investigate and possibly ameliorate this issue. Additionally, investigating the usefulness of AES feedback in helping ESL students improve their writing might also help instructors to utilize such assignments for formative feedback. Third, our analysis was specifically limited to one AES system within one course; additional testing should be conducted to examine the generalizability and applicability of our findings and suggestions.

Finally, methodological issues related to scale truncation might account for some of the significant differences found between instructor and AES scoring consistency. For the instructor-AES agreement analyses, coefficients are likely to be conservative in their agreement estimation due to the heavily skewed score distributions. In other words, these estimates may be lower than would be expected due to the penalties they assign for a large number of scores falling in a score category (Feinstein & Cicchetti, 1990). The majority of students in our sample received high scores and this ceiling effect is likely to have contributed to the great discrepancy between percent agreement and the chance-corrected agreement indices. Future research should be conducted using a rubric with more scale points, in order to better differentiate between student ability levels and better conduct comparability analyses.

Conclusion

Despite study limitations, our findings suggest that ESL students were consistently given lower scores between the AES and instructor grader. This finding suggest that future research may consider incorporating DIF analysis into AES systems in order to decipher whether these differences are a product of differential item functioning or impact. Though we were unable to conduct this analysis due to a low number of participants for the instructor-grading sample, we would encourage future researchers investigating language effects and AES use in MOOCs to utilize this methodology. Overall, our findings revealed that both the AES and instructor-graded non-native English speakers lower, and that instructor and AES score comparability was better for native English speakers. This research suggests the need to further evaluate the use of AES graders in MOOCs for non-native English speakers. These systems might be better utilized as formative, low-stakes assessments or to help students reach goals around English literacy. Additionally, future research may need to qualitatively evaluate the use of AES systems for non-native English speakers, in terms of usefulness of feedback, navigability, and perceptions of assignment fairness.

Acknowledgement

This study was supported by the MOOC Research Initiative (MRI), funded by the Bill & Melinda Gates Foundation.

References

Attali, Y., & Burstein, J. (2006). Automated essay scoring with e-rater® v.2. *Journal of Technology, Learning, and Assessment, 4*(3), 1-31.

Balfour, S. P. (2013). Assessing writing in MOOCs: Automated essay scoring and calibrated peer review™. *Research & Practice in Assessment, Special Issue: MOOCs & Technology, 8,* 40-48.

Burstein, J., & Chodorow, M. (1999, June). Automated essay scoring for nonnative English speakers. In *Proceedings of the ACL99 workshop on computer-mediated language assessment and evaluation of natural language processing.* College Park, MD.

Byerly, A. (2012). Before you jump on the bandwagon…*Chronicle of Higher Education, 59*(2), 34.

Casanave, C. P., & Hubbard, P. (1992). The writing assignments and writing problems of doctoral students: Faculty perceptions, pedagogical issues, and needed research. *English for Specific Purposes, 11*(1), 33-49.

Chen, C-F. E., & Cheng, W-Y. E. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning & Technology, 12*(2), 94-112.

Clifford, J., Merschel, L., & Munné, J. (2013). Surveying the Landscape: What is the Role of Machine Translation in Language Learning?. @ tic.revista d'innovació educativa, 10, 108-121.

Condon, W. (2013). Large-scale assessment, locally-developed measures, and automated scoring of essays: Fishing for red herrings. *Assessing Writing, 18*(1), 100-108.

Cushing Weigle, S. (2013). English language learners and automated scoring of essays: Critical considerations. *Assessing Writing, 18*, 85-99.

Dikli, S. (2006). An overview of automated scoring of essays. *The Journal of Technology, Learning, and Assessment, 5*(1), 1-36.

Dikli, S. & Bleyle, S. (2014). Automated essay scoring feedback for second language writers: How does it compare to instructor feedback?. *Assessing Writing, 22*, 1-17.

Dolan, V. L. B. (2014). Massive online obsessive compulsion: What are they saying out there about the latest phenomenon in higher education? *The International Review of Research in Open and Distance Learning, 15*(2), 268-281.

Goodfellow, R. & Lea, M.R. (2005). Supporting writing for assessment in online learning. *Assessment & Evaluation in Higher Education, 30*(3), 261 – 271.

Guo, F. (2009). Fairness of automated essay scoring of GMAT® AWA. Graduate Management Admission Council® Research Report. Retrieved from http://www.gmac.com/~/media/files/gmac/research/research-report-series/rr0901_awafairness.pdf

Feinstein, A.R., & Cicchetti, D.V. (1990). High agreement but low kappa: The problems of two paradoxes. *Journal of Clinical Epidemiology, 43*(6), 543-549.

Furnborough, C., & Truman, M. (2009). How do adult beginner distance language learners respond to assignment feedback? A qualitative study. *Distance Education, 30*(3), 399-418.

Hamel, R.E. (2007). The dominance of English in the international scientific periodical literature and the future of language use in science. *AILA Review 20*, 53-71.

Hollands, F., & Tirthali, D. (2014). Why do Institutions Offer MOOCs?. *Online Learning: Official Journal of the Online Learning Consortium, 18*(3). Retrieved from http://olj.onlinelearningconsortium.org/index.php/jaln/article/view/464

Karabulut, A., Levelle, K., Li, J., & Suvorov, R. (2012). Technology for French learning: A mismatch between expectations and reality. *CALICO Journal, 29*(2), 341–366.

Kim, J. (2012, March). Why every university does not need a MOOC. *Inside Higher Ed*. Retrieved from https://www.insidehighered.com/blogs/technology-and-learning/why-every-university-does-not-need-mooc

Larson-Guenette, J. (2013), “It's just reflex now”: German language learners’ use of online resources. *Die Unterrichtspraxis/Teaching German, 46,* 62–74.

Liyanagunawardena, T. R., Adams, A. A., & Williams, S. A. (2013). MOOCs: A systematic study of the published literature 2008-2012. *The International Review of Research in Open and Distance Learning, 14*(3), 202-227.

Liu, X., Liu, S., Lee, S.-h., & Magjuka, R. J. (2010). Cultural differences in online learning: International student perceptions. *Educational Technology & Society, 13*(3), 177-188.

McCurry, D. (2010). Can machine scoring deal with broad and open writing tests as well as human readers? *Assessing Writing, 15*(2), 118-129.

Methitham, P. (2009). *An Exploration of culturally-based assumptions guiding ELT practice in Thailand, a non-colonized nation*. Doctoral dissertation, Indiana University of Pennsylvania.

Meyer, J. P. & Zhu, S. (2013). Fair and equitable measurement of student learning in MOOCs: An introduction to item response theory, scale linking, and score equating. *Research & Practice in Assessment Special Issue, 8*, 26-39.

Morgan, T., & Carey, S. (2009). From open content to open course models: increasing access and enabling global participation in higher education. *The International Review of Research in Open and Distance Learning, 10*(5).

Pennycook, A. (1999). Introduction: Critical approaches to TESOL. *TESOL Quarterly, 33*(3), 329-348.

Peters, O. (2002). Distance education in transition: New trends and challenges. Oldenburg: Bibliotheks- und Informations system der Universität Oldenburg.

Phillipson, R. (1992). *Linguistic imperialism.* Oxford: Oxford University Press.

Reilly, E. D., Stafford, R. E., Williams, K. M., & Corliss, S. B. (2014). Evaluating the validity and applicability of automated essay scoring in two massive open online courses. *The International Review of Research in Open and Distributed Learning*, *15*(5).

Sandeen, C. (2013). Assessment’s place in the new MOOC world. *Research & Practice in Assessment, Special Issue: MOOCs & Technology, 8*, 5-12.

Tan, F. (2009). Tri-fold transformation: An international adult students’ reflections on online learning. *Adult Learning, (20)*3-4, 38-40.

Uzuner, S. (2009). Questions of culture in distance learning: A research review. *The International Review of Research in Open and Distance Learning*, 10(3).

Yuan, L., & Powell, S. (2013). MOOC and open education: Implications for higher education. Centre for

Educational Technology & Interoperability Standards (white paper). Retrieved from http://publications.cetis.ac.uk/2013/667

Young, S. L (2013). MOOCS: Revolutionizing education or the latest business opportunity?. *Fukuoka University Review of Commercial Sciences, 58*(1-2), 173-187.

Walkow, J. C., & Reilly, E. (2014). Are we ready for robots to grade? *The Chronicle of Higher Education, 61.* Retrieved from http://chronicle.com/article/Are-We-Ready-for-Robots-to/148723/

Watters, A. (2013). MOOC mania: Debunking the hype around massive open online courses. *The Digital Shift.* Retrieved from: http://www.thedigitalshift.com/2013/04/featured/got-mooc-massive-open-online-courses-are-poised-to-change-the-face-of-education/

Williams, L. (2006). Web-Based Machine Translation as a Tool for Promoting Electronic Literacy and Language Awareness. *Foreign Language Annals, 39*: 565–578.

The Role of Social Influence in Anxiety and the Imposter Phenomenon

Christy B. Fraenza
Walden University

Abstract
High anxiety levels have been associated with high levels of the imposter phenomenon (IP), a negative experience of feeling like a fraud. This study was designed to explore IP among graduate students and to determine whether a difference exists between online graduate students and traditional graduate students. The theoretical foundation of this study was social influence, which holds that students may feel pressured in a traditional setting because of the social cues of peers and instructors, as well as institutional norms. This quantitative study used a between-subjects design to compare 2 independent samples (115 online students & 105 traditional students). The study used a cross-sectional survey design, with 4 different measures: the Clance Imposter Phenomenon Scale, the Zung Self-Rating Anxiety Scale, the Perfectionistic Self-Presentation Scale, and a basic demographic survey. Results indicated that traditional graduate students had significantly higher IP scores than online graduate students. Results also indicated a significant, positive relationship between IP scores and anxiety scores. Regression analysis indicated that perfectionism was the most influential predictor of IP scores, followed by anxiety and program type. Because the scale used in this study explored socially prescribed perfectionism, the results appear to suggest an underlying social component to IP.

Introduction

The invention of the Internet has changed how individuals interact with one another on a daily basis. People use chatrooms, e-mails, and social networking websites instead of phone calls and lunch meetings to keep in touch. The advent of the Internet has also paved the way for online educational programs. Many of these programs utilize e-mail and asynchronous discussions to submit work and interact with peers and instructors. This type of interaction, typically referred to as computer-mediated communication (CMC), has changed the college classroom. In this setting, many of the social cues that may influence individuals are lost, such as body language. As outlined by Salter (2003), this type of interaction allows students to develop their own opinions over time rather than quickly responding to questions in a physical classroom. For those individuals who experience high levels of anxiety, the absence of these cues and pressures to respond may provide a level of comfort.

Literature Review

Individuals who experience the imposter phenomenon (IP) are stuck in a negative cycle of emotions. While they are typically very successful, they are unable to fully experience and own their achievements. They also may avoid extra academic pursuits, such as participating in conferences (Harvey, 1982). Regardless of success, these individuals believe they have somehow fooled everyone into believing they are more competent than they really are. These individuals live in a negative cycle of fear of discovery, over preparation, and experience high levels of anxiety (Kolligan & Sternberg, 1991).

The Imposter Phenomenon and Anxiety

Researchers have indicated that IP is positively related to anxiety (Chae, Piedmont, & Estadt, 1995). As discussed by Kolligan and Sternberg (1991), a key feature of IP is social anxiety, especially when there is imminent evaluation. The fear of discovery fuels anxiety levels, which may contribute to high levels of perfectionism as well as the *imposter cycle*. As outlined by Caselman et al., (2006), this cycle is characterized by a fear of failure that often leads to perfectionism and over preparation. This fear of discovery often pushes the individual to intensely focus on preparation in order to succeed at a task. However, this negative cycle of anxiety is often followed by success, which only reinforces the belief that their success was not legitimate (Caselman et al., 2006), and so the cycle continues.

Graduate Students

Researchers have explored IP in professions where anxiety is likely present, including family medicine residents (Oriel, Plane, & Mundt, 2004), medical students (Henning, Ey, & Shaw, 1998), physician assistants (Prata & Gietzen, 2007), and marketing managers (Fried-Buchalter, 1992; 1997). Little research has explored IP in the population of graduate students, but it is commonly known that graduate school is full of anxiety-provoking assignments and experiences (Hadijoannou, Shelton, & Dhanarattigannon, 2002). Horne (2011) qualitatively explored IP among a small sample of doctoral students and found that they commonly experienced anxiety and felt that they had been admitted into their programs by mistake.

Literature on IP is limited regarding graduate students, and no known studies have compared online graduate students to traditional graduate students. Previous studies on graduate students have indicated that online students experience lower levels of anxiety (DeVaney, 2010). Researchers have theorized that the loss of social cues and pressures in electronic communications may reduce anxiety associated with asking for help (Kitsantas & Chow, 2007) and may create a more comfortable, open environment where all members are equal (Sullivan, 2002). Many cues about the context of an interaction are not present in electronic communication, such as body language, nonverbal cues, physical appearance, and emotional reactions (Parks & Floyd, 1996). Without this information, communication online should result in less social influence and conformity in comparison to face-to-face communications (Parks & Floyd, 1996). Because face-to-face communications are absent, or limited, in online graduate programs, it was theorized that online graduate students would experience less anxiety and less intense IP, when compared to their traditional counterparts.

Social Influence

The role of anxiety in IP was at the foundation of this study, as well as the potential impact of social influences. As discussed by Cialdini and Goldstein (2004), social influence refers to a change in behaviors or beliefs in response to experienced social forces. As discussed by Liao and Hsieh (2011), “another term for social influence is peer group pressure, or the pressure on a person to conform to a distinct group resulting in a specific behavior” (p. 888). Individuals in a traditional classroom setting may experience increased levels of social influence and pressure due to participation requirements within the classroom. Individuals with underlying anxiety may experience increased anxiety in this setting due to speaking in class and frequent interpersonal contact (Ioakimidis, 2010). However, this added pressure and anxiety may not be present in the online classroom as individuals are able to have more control over their communication with others (Okdie, Guadagno, Bernieri, Geers, & Mclarney-Vesotski, 2011).

Social influences impact individuals every day during regular activities, including academic studies. Students are constantly confronted with social interactions when others around them attempt to change their attitudes, behavior, or beliefs (Mugny & Guimond, 2007). Students experiencing high levels of anxiety may feel pressured in a traditional setting because of the perceived social cues of peers and instructors, as well as the pressures to conform to institutional norms. Such pressure in the social setting may limit a student’s participation and learning and this added anxiety may be a contributing factor to the experience of IP. However, research has indicated that students feel more able to be themselves in an online classroom setting (Sullivan, 2002). Students have also stated that they felt the online classroom allowed everyone to be equal, as nobody knew what others looked like (Sullivan, 2002). Research has also indicated that students feel more comfortable asking for help through e-mail than in person (Kitsantas & Chow, 2007). Researchers have theorized that this may be due to students’ ability to construct a question and because online communication frees them from the nonverbal cues and social influence of the instructor and classmates (Kitsantas & Chow, 2007). This research was designed to build on the body of literature about online learning and CMC by exploring anxiety levels and IP levels within this setting.

Summary

Individuals dealing with IP live their lives in constant fear of being discovered as less intelligent or competent than others believe them to be (Clance & Imes, 1978). Even in the face of frequent success, this underlying fear of discovery pushes them to reach toward impossibly high standards of achievement and success. Instead of attributing their achievements to their abilities, they often attribute their success to luck, hard work, or the mistakes of others (Clance & O’Toole, 1988). The negative cycle of IP leaves them in a state of uncomfortable anxiety, especially in the face of imminent evaluations (Kolligan & Sternberg, 1991).

Frequent evaluations are a consistent thread across graduate programs; students are burdened by high workloads and deadlines. For an individual experiencing IP, a graduate school program is an anxiety-ridden chapter in their lives. It was theorized that individuals in online graduate programs may experience less underlying anxiety due to less social influence and interaction.

Method

Procedures

Written information introducing the study was posted to the participant pool service of a large online university. Enrolled online graduate students who had access to the participant pool were invited to participate. In addition, an announcement regarding the study was posted on the LinkedIn group page for the online university. Traditional graduate students were sent the same information via a listserv for the graduate student association at the large university. Graduate students that signed up for the listserv would receive announcements of interest to the graduate student population. But because the listserv is voluntary and not all graduate students would be on the list, it was decided to also contact each department head directly, who then forwarded information about the study to graduate students enrolled in their own departments. Students were invited to participate in the study by using the link within an e-mail that led them to the informed consent and questionnaires. After completing the surveys, participants were thanked for their participation and debriefed.

Research Questions & Hypotheses

Research Question 1: Is there a significant difference in IP scores, as measured by the Clance Imposter Phenomenon Scale (CIPS), between online graduate students and traditional graduate students?

Null Hypothesis ($H_0 1$): There will not be a statistically significant difference between the IP scores of online graduate students and traditional graduate students, as measured by the CIPS.

Alternate Hypothesis ($H_1 1$): There will be a statistically significant difference in the IP scores, as measured by the CIPS, of online graduate students when compared to the IP scores of traditional graduate students. It is hypothesized that online graduate students will have lower IP scores than traditional graduate students.

Research Question 2: What is the nature of the relationship between IP, as measured by CIPS, and anxiety scores, as measured by the Zung Self-Rating Anxiety Scale (SAS)?

Null Hypothesis ($H_0 2$): There will be no statistically significant relationship between IP, as measured by the CIPS, and anxiety, as measured by the SAS.

Alternate Hypothesis ($H_1 2$): There will be a significant, positive relationship between IP, as measured by the CIPS, and anxiety, as measured by the SAS (Bernard, 2002; Chae et al., 1995).

Research Question 3: Is there a difference in anxiety scores, as measured by the SAS, between online graduate students and traditional graduate students?

Null Hypothesis ($H_0 3$): There will not be a statistically significant difference between the anxiety scores, as measured by the SAS, of online graduate students and traditional graduate students.

Alternate Hypothesis ($H_1 3$): There will be a statistically significant difference in the anxiety scores, as measured by the SAS, of online graduate students when compared to traditional graduate students. It is hypothesized that online graduate students will have significantly lower anxiety scores.

Participants

Participants were 220 graduate students from two different universities, with 105 enrolled in traditional programs and 115 enrolled in online programs. The sample was 76% White and 16.4% Black. Fifty-four participants were enrolled in a master's level program and 166 were pursuing a doctoral degree. Participants came from a variety of majors: 47.7% social science, 15% science, 5.9% business, 2.3% technology, 2.3% medicine, 8.2% education, 2.3% arts, and 16.4% other.

Measures

Demographics. A demographic questionnaire collected basic demographic information from each participant. Information included on this questionnaire consisted of: age, ethnicity, educational program, year in educational program, and previous graduate school experience.

Clance Imposter Phenomenon Scale (CIPS). The CIPS consists of 20 items that assess various aspects of IP, using a Likert-scale format (Clance, 1985). Participants rate their agreement or disagreement using a 5-point scale (1 = *not at all true*, 2 = *rarely*, 3 = *sometimes*, 4 = *often*, 5 = *very true*). Total scores can therefore range from 20 to 100, with higher scores corresponding to more severe IP.

Holmes, Kertay, Adamson, Holland, and Clance (1993) found the CIPS had high internal consistency (coefficient alpha of .96). In comparison to the other used IP scale, the Harvey IP Scale, the CIPS has been shown to be a more reliable and sensitive instrument (Holmes et al., 1993).

Zung Self-Rating Anxiety Scale (SAS). The SAS is a 20-item questionnaire in which the participant is asked to indicate how often they have experienced certain symptoms associated with anxiety, using a Likert-scale format (1 = *none OR a little of the time*, 2 = *some of the time*, 3 = *good part of the time*, 4 = *most OR all of the time*) (Zung, 1971). Participants report the frequency of symptoms over the previous week. The 20-items include both psychological symptoms, such as fear and nervousness, as well as physiological symptoms, such as sweating and insomnia (Bitsika, Sharpley, & Bell, 2009). Total scores can range from 20 to 80, with a score above 36 indicating a level of anxiety that requires further assessment or treatment (Sharpley, Bitsika, & Christie, 2010). De la Ossa, Martinez, Herazo, and Campo (2009), determined that the SAS had acceptable internal consistency with a coefficient alpha of .77. Sharpley and Rogers (1985) reported a coefficient alpha of .79.

Perfectionistic Self-Presentation Scale (PSPS). The PSPS was developed to assess an individual's level of focus on appearing perfect to other people, as well as an individual's level of avoidance regarding outward expressions of their flaws (Hewitt et al., 2003). The PSPS is a 27-item questionnaire, in which participants are asked to rate their level of agreement, using a Likert-scale format (1 = *disagree strongly*; 4 = *neutral/undecided*; and 7 = *strongly agree*). The PSP includes three different dimensions of perfectionistic self-presentation, including nondisclosure of imperfection, nondisplay of imperfection, and perfectionistic self-presentation. Reynolds, Dowell, Peeters, Beene, Espino, and Longstreth (2007) described nondisclosure of imperfection as an individual's openness to reveal their own perceived flaws. Non-display of imperfection was described as an individual's focus on hiding any perceived imperfections and perfectionistic self-presentation was described as actively promoting one's self in order to have the appearance of being perfect (Reynolds et al., 2007). Three-week test-retest reliability values have indicated a high degree of consistency (perfectionistic self-promotion = .83; non-display of imperfection = .84; nondisclosure of imperfection = .74). Results have also suggested that these dimensions are linked with measures of social anxiety, in particular, the non-display of imperfection dimension (Hewitt et al., 2003). This measure was included to capture the impact of the social environment on the variables included within this study.

Results

Research Question 1

Research question 1 asked if there was a significant difference in IP scores, as measured by the Clance Imposter Phenomenon Scale (CIPS), between online graduate students and traditional students. In order to determine if there was a significant difference between online graduate students and traditional graduate students in IP scores, an independent *t*-test was used. Results of a Levene's Test for Equality of Variances indicated a sig value of .559, which indicated that the variances were equal. Results indicated a statistically significant difference between traditional graduate students and online graduate students on the CIPS, t (193) = 4.948, $p < .001$. Traditional graduate students had a higher mean CIPS score of 65.68 (SD = 15.59) compared to the mean online graduate CIPS score of 54.17 (SD = 16.85). The effect size for this analysis (d = .71) indicated a medium effect size. These results are provided in Table 1.

Research Question 2

Research question 2 asked what the nature of the relationship was between IP, as measured by CIPS, and anxiety scores, as measured by the SAS. A Pearson correlation was used to determine if there was a significant relationship between IP and anxiety. Results indicated that there was a significant, positive relationship between CIPS scores and SAS scores, r (187) = .405, p < .001. These results indicated a medium effect size.

Research Question 3

Research question 3 asked if there was a difference in anxiety scores, as measured by the SAS, between online graduate students and traditional graduate students. In order to determine if there was a significant difference between online graduate students and traditional graduates in anxiety scores, an independent t-test was used. The results of a Levene's Test for Equality of Variance indicated a sig value of .451, which indicated that the variances were equal. Traditional graduate students had higher anxiety scores (M = 37.49, SD = 6.01) than online graduate students (M = 36.88, SD = 6.76), however this difference was not significant, t (211) = .693, p = .489. The effect size for this analysis (d = .10) indicated a small effect. These results are provided in Table 1.

Table 1
CIPS Scores and SAS Scores for Online and Traditional Graduate Students

	Program Type			
	Online	Traditional	*t*	*df*
CIPS Scores	54.17 (16.85)	65.68 (15.59)	4.948*	193
SAS Scores	36.88 (6.76)	37.49 (6.01)	.693	211

Note. * = $p \leq .001$. Standard deviation appears in parentheses below means.

Additional Analysis

Standard multiple linear regression was carried out to determine the effect of perfectionism, anxiety, and program type on IP scores. This was a statistically significant model, F (3, 172) = 84.72, p < .001, indicating these results were unlikely to have been obtained by chance. The adjusted R^2 indicated that 58.9% of the variance in IP scores can be explained by variances in the three predictor variables. The analysis suggested that perfectionism (β = .638) was the most influential predictor, followed by anxiety (β = .192), and program type (β = - .157). All three predictor variables were shown to be statistically significant predictors of IP score; these results are provided in Table 2.

Table 2
Regression Model with IP score as the Criterion Variable

	Unstandardized Coefficients	Standardized Coefficients	*t*	*Sig.*
Constant	-.343			
PSPS_Scores	.448	.638	12.100	.000
SAS_Scores	.543	.192	3.744	.000
Program Type	-5.403	-.157	-3.157	.002

Note: Adjusted R^2 = 58.9%; F (3, 172) = 84.72, p < .001

Discussion

Results supported the hypothesis that online graduate students would show significantly lower IP scores than traditional graduate students, as well as the hypothesis that there would be a significant, positive relationship between IP scores and anxiety scores. Results indicated that traditional students had higher anxiety scores than online graduate students, however this result was not significant. Additionally, perfectionism was identified as the most influential predictor of IP scores.

Results supported the premise that IP would be present within a sample of graduate students. Traditional graduate student participants had a mean score that indicated frequent imposter feelings (M = 65.68, SD = 15.59). Online graduate student participants had a mean score that indicated a moderate level of IP (M = 54.17, SD = 16.85). Results also showed high levels of anxiety for both traditional graduate students (M = 37.49, SD = 6.01) and online graduate students (M = 36.88, SD = 6.76). A score of 36 or over on the SAS is considered high; in fact, it would suggest that an individual might need further assessment (Sharpley et al., 2010). Overall, these results indicate that graduate students, whether they are enrolled in online programs or traditional programs, are dealing with anxiety and IP.

Results indicated a significant, positive relationship between anxiety scores and IP scores. This finding supports previous research that IP and anxiety are positively related and adds to the body of literature that has explored this correlation (see Bernard, 2002; Chae et al., 1995; Ives, 2010). It also appears to further support the assumption that individuals in high-anxiety provoking professions or programs are likely to experience IP.

Although results indicated that traditional graduate students had significantly higher IP scores when compared to online graduate students, a significant difference in anxiety scores was not found. This finding seems to suggest that there must be another component involved in IP scores among graduate students. Following further analyses, perfectionism was found to be the most significant predictor of IP scores. The PSPS was used in order to assess the impact of the social environment on the variables within this study, as it was designed to explore a person's level of focus on appearing perfect to other people, as well as an individual's level of avoidance regarding outward expressions of their flaws (Hewitt et al., 2010). Socially prescribed perfectionism may be influential in interpersonal situations, because it has been shown to increase self-appraisal and attention, which in turn influences an individual's perceptions of their own abilities and behaviors (Aldan & Bieling, 1993). As stated by Ives (2010), the IP experience is associated with worries about being discovered as a fraud, feelings of having tricked others, and a tendency to attribute success to outside influences. These factors of IP all seem to suggest an underlying social component to the IP experience; the individual with IP feels they have conned others about their abilities and are fearful of being discovered for doing so. Ultimately, perfectionism was found to be a powerful predictor of IP scores, even more so than type of program and anxiety, which seems to suggest that the social environment played a role in IP scores. As stated by Burnkrant and Cousineau (1975), one of the greatest influences on an individual's behavior is simply the influence of others around them. Because the scale used to assess perfectionism was only focused on socially prescribed perfectionism that is ultimately driven by social forces, it seems that an underlying social component could be a factor in the results. According to Hewitt (as cited by Benson, 2003), "interpersonal needs are what drive the perfectionistic behavior" (p. 3). These needs are the focus of the PSPS, which includes such items as "If I seem perfect, others will see me more positively" and "Errors are much worse if they are made in public rather than in private" (Hewitt et al., 2003, p. 1). However, it is also possible that personality characteristics not explored within this study played a role in these results. Researchers have explored personality characteristics related to IP and found positive correlations between neuroticism and IP (Bernard et al., 2002; Chae et al., 1995), however this was not explored in this study. As such, it is important to recognize that the results of this study do have limitations.

Limitations of the Study

One limitation of this study is the inability to infer any type of cause and effect relationship between the variables of interest. The design of the study only allowed for a numerical picture of the trends and relationships among the variables of interest. Because there was no intervention or manipulation of variables, it is beyond the scope of this study to determine if a particular learning environment caused certain levels of anxiety and IP. While significant differences were found between these groups on IP scores, it cannot be concluded that traditional graduate programs cause IP.

Another concern is that the results are not generalizable to the greater graduate student population, as the sample utilized for this study was predominantly White. As described by the U.S. Department of Education, National Center for Education Statistics (2012), from 1976 to 2010, the percentage of White students fell from 83% to only 61%, while the percentage of Black students rose from 9% to 14% and the percentage of Asian/Pacific Islander students rose from 2% to 6%. Given this information regarding the larger graduate student population, the sample utilized for this study is not a true representation of the graduate student population.

Another limitation is that it is impossible to determine if the groups used in this study were equal on other variables not measured. For instance, one specific issue not addressed by this study was the timing of the questionnaires. Ioakimidis (2007) found that online students' computer anxiety tends to decrease over time, while the anxiety of traditional students remains steady. Because of the differences in the two different academic institutions within this study, it was not possible to ensure that the questionnaires were distributed at the same time of the academic quarter or semester. Therefore, this potential effect could not be explored.

Another factor not directly assessed in this study were differences in attrition between online graduate students and traditional graduate students. Patterson and McFadden (2009) found that students in an online Master's of Business Administration (MBA) program and a Master's in Communication Sciences and Disorders (MCSD) were more likely to drop out than students enrolled in the traditional version of these programs. It is unclear what the attrition rates are between online students and traditional students on a larger scale, but if there are differences, this is a limitation to this study. It is possible that online students could drop out at higher rates (possibly due to stress and anxiety) when compared to traditional students. If students who have experienced high levels of stress dropped out of their programs, they were not included within this study.

One last limitation related to the sample is that it was not random. Rather, the sample consisted of only graduate students who volunteered to complete the survey. It is very possible that those that volunteered to participate are somehow different from those that did not participate. For instance, those students who participated may have fewer responsibilities that allowed them the extra time to complete the surveys. Furthermore, researchers have noted that anxiety among imposters tends to increase during times of evaluation (Leary et al., 2000) and this could have also impacted participation. For instance, some students may have been completing final steps in their programs; this added anxiety related to evaluations could have kept from them participating.

Future Directions

This study has contributed to the overall literature on IP and has explored whether IP is experienced differently between learning environments. Because both traditional graduate students and online graduate students showed high levels of anxiety and IP, it is important to develop further research aimed at reducing these experiences within the population of graduate students. Ives (2010) investigated the impact of an orientation course on IP and anxiety among a sample of online graduate students and found that such a course could reduce the IP experience over time. However, this orientation course did not provide relief of anxiety (Ives, 2010). Further research into such interventions for both online graduate students and traditional graduate students is needed in order to reduce the negative experiences associated with IP and anxiety.

Furthermore, more research is needed to explore the relationship among social influence, anxiety, perfectionism, and IP. Future research could benefit from more in-depth assessments utilizing qualitative components in order to better understand the complex interactions among these variables. In addition, it is still unclear how the IP experience impacts academic performance. Future research could explore how the

IP experience impacts the academic experience of graduate students and whether this experience is different between online and traditional educational settings.

Conclusion

This study contributes to the literature focused on IP by exploring differences between online graduate students and traditional graduate students. Online graduate students had lower IP scores when compared to traditional graduate students and socially prescribed perfectionism was identified as the strongest predictor of IP scores. Because socially prescribed perfectionism involves a social component, these results suggest that social influences may play a role in the IP experience.

References

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs: General And Applied*, *70*(9), 1-70. doi:10.1037/h0093718

Alden, L.E., & Bieling, P.M. (1993, August). *Perfectionism in an interpersonal context*. A paper presented at the Annual Meeting of the American Psychological Association, Ontario, Canada.

Austin, J., Cameron, T., Glass, M., Kosko, K., Marsh, F., Abdelmagid, R., & Burge, P. (2009). First semester experiences of professionals transitioning to full-time doctoral study. *The College Student Affairs Journal*, *27*(2), 194-214.

Bandura, A., Ross, D., & Ross, S. A. (1963). A comparative test of the status envy, social power, and secondary reinforcement theories of identificatory learning. *Journal of Abnormal & Social Psychology*, *67*(6), 527-534.

Baralt, M., & Gurzynski-Weiss, L. (2011). Comparing learners' state anxiety during task-based interaction in computer-mediated and face-to-face communication. *Language Teaching Research*, *15*(2), 201-229. doi:10.1177/0265532210388717

Beck, A., Epstein, N., Brown, G., & Steer, R. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology, 56*(6), 893-897. doi:10.1037/0022-006X.56.6.893

Benson, E. (2003). The many faces of perfectionism. *Monitor*, *34*(10), 18. Retrieved from http://www.apa.org/monitor/nov03/manyfaces.aspx

Bernard, N.S., Dollinger, S.J., & Ramaniah, N.V. (2002). Applying the big five personality factors to the imposter phenomenon. *Journal of Personality Assessment, 78*(2), 321-333.

Bitsika, V., Sharpley, C. F., & Bell, R. (2009). The contribution of anxiety and depression to fatigue among a sample of Australian university students: Suggestions for university counsellors. *Counselling Psychology Quarterly*, *22*(2), 243-255. doi:10.1080/09515070903216929

Burkholder, G. (n.d.). Study notes: Sample size analysis for quantitative studies. *Research Methods.* Los Angeles, CA: Laureate Education, Inc.

Burnkrant, R.E., & Cousineau, A. (1975). Informational and normative social influence in buyer behavior. *Journal of Consumer Research, 2*(12), 206-215.

Byrnes, K., & Lester, D. (1995). The imposter phenomenon in teachers and accountants [Abstract]. *Psychological Reports, 77*(1), 350.

Caselman, T., Self, P., & Self, A. (2006). Adolescent attributes contributing to the imposter phenomenon. *Journal of Adolescence, 29*(3), 395-405. doi:10.1016/j.adolescence.2005.07.003

Castro, D., Jones, R., & Mirsalimi, H. (2004). Parentification and the impostor phenomenon: An empirical investigation. *American Journal of Family Therapy, 32*(3), 205-216. doi:10.1080/01926180490425676

Chae, J.H., Piedmont, R.L., Estadt, B.K., & Wicks, R.J. (1995). Personological evaluation of Clance's Imposter Phenomenon Scale in a Korean sample. *Journal of Personality Assessment, 65*(3), 468-485.

Chrisman, S.M, Pieper, W.A., Clance, P.R., Holland, C.L., & Glickauf-Hughes, C. (1995). Validation of the Clance Imposter Phenomenon Scale. *Journal of Personality Assessment, 65*(3), 456-467.

Cialdini, R. & Goldstein, N. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, *55*, 591-621.

Cialdini, R.B., & Trost, M.R. (1998). Social influence: Social norms, conformity, and compliance. In D.T. Gilbert, S.T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology*, (4th ed., Vol. 2, pp. 151-193). New York, NY: McGraw Hill.

Cialdini, R.B., Vincent, J.E., Lewis, S.K., Catalan, J., Wheeler, D., & Darby, B.L. (1975). Reciprocal concessions procedure for inducing compliance: The door-in-the-face technique. *Journal of Personality and Social Psychology*, *31*(2), 206-215.

Claidière, N., & Whiten, A. (2012). Integrating the study of conformity and culture in humans and nonhuman animals. *Psychological Bulletin*, *138*(1), 126-145. doi:10.1037/a0025868

Clance, P.R. & Imes, S.A. (1978). The imposter phenomenon in high achieving women: Dynamics and therapeutic intervention. *Psychotherapy: Theory, Research and Practice, 15*(3), 241-247.

Clance, P.R. (1985). *The Imposter Phenomenon: When success makes you feel like a fake.* New York, NY: Bantam Books.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative and mixed methods approaches* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.

Cromwell, B., Brown, N., Sanchez-Huceles, J., & Adair, F. (1990). The imposter phenomenon and personality characteristics of high school honor students. *Journal of Social Behavior and Personality, 5*(6), 563-573.

de la Ossa, S., Martinez, Y., Herazo, E., & Campo, A. (2009). Study of internal consistency and factor structure of three versions of the Zung's rating instrument for anxiety disorders. *Colombia Médica*, *40*(1), 78-84.

Deutsch, M., & Gerard, H.B. (1955). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal and Social Psychology, 51*, 629-636.

Dolinski, D. (2011). A rock or a hard place: The foot-in-the-face technique for inducing compliance without pressure. *Journal of Applied Social Psychology*, *41*, 1514-1537.

Flett, G.L., Hewitt, P.L., Blankstein, K.R., & Gray, L. (1998). Psychological distress and the frequency of perfectionistic thinking. *Journal of Personality and Social Psychology, 75*(5), 1363-1381.

Flett, G.L., Hewitt, P.L., Demerjian, A., Sturnam, E.D., Sherry, S.B., & Cheng, W. (2011). Perfectionistic automatic thoughts and psychological distress in adolescents: An analysis of the perfectionism cognitions inventory. *Journal of Rational-Emotive Cognitive-Behavior Therapy*. doi: 10.1007/s10942-011-0131-7

Fogg, P. (2009). Grad-school blues. *Chronicle of Higher Education*, *55*(24), 12-16.

Freedman, J. L., & Fraser, S. C. (1966). Compliance without pressure: The foot-in-the-door technique. *Journal of Personality & Social Psychology*, *4*(2), 195-202.

Fried-Buchalter, S. (1997). Fear of success, fear of failure, and the imposter phenomenon among male and female marketing managers. *Sex Roles, 37*(11), 847-859. doi:10.1007/BF02936343

Gliem, J.A. & Gliem, R.R. (2003). Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. *2003 Midwest Research to Practice Conference in Adult, Continuing, and Community Education*. Retrieved from http://hdl.handle.net/1805/344

Goethals, G.R. (2003). A century of social psychology: Individuals, ideas, and investigations. In M.A. Hogg, & J. Cooper (Eds.), *The Sage handbook of social psychology* (pp. 4-24). Thousand Oaks, CA: SAGE. Doi: 10.4135/978-1-84860-822-1.n1

Gravois, J. (2007). You're not fooling anyone. *Chronicle of Higher Education, 54*(11).

Green, K.E. (1997). Psychosocial factors affecting dissertation completion. *New Directions for High Education*, *99*, 57-64.

Guéguen, N. (2003). Fund-raising on the web: The effect of an electronic door-in-the-face technique on compliance to a request. *CyberPsychology & Behavior*, *6*(2), 189-193.

Hadjioannou, X., Shelton, N., Fu, D., & Dhanarattigannon, J. (2007). The road to a doctoral degree: Co-travelers through a perilous passage. *College Student Journal*, *41*(1), 160-177.

Harvey, J.C. (1981). The imposter phenomenon and achievement: A failure to internalize success. Ph.D. dissertation, Temple University. Retrieved from Dissertations & Theses: Full Text. (Publication No. AAT 8210500).

Harvey, J.C. & Katz, C. (1985). *If I'm so successful why do I feel like a fake? The imposter phenomenon*. New York, NY: Simon and Shuster.

Heiman, G.W. (2003). *Basic statistics for the behavioral sciences* (4th ed.). Boston, MA: Houghton Mifflin Co.

Henning, K., Ey, S., & Shaw, D. (1998). Perfectionism, the imposter phenomenon and psychological adjustment in medical, dental, nursing and pharmacy students. *Medical Education, 32*, 456-464.

Hewitt, P.L., Flett, G.L., Sherry, S.B., Habke, M., Parkin, M., Lam, R.W., & Stein, M.B. (2003). The interpersonal expression of perfection: Perfectionistic self-presentation and psychological distress. *Journal of Personality and Social Psychology*, *84*, 1303-1325.

Holmes, S., Kertay, L., Adamson, L., & Holland, C. (1993). Measuring the imposter phenomenon: A comparison of Clance's IP scale and Harvey's I-P scale. *Journal of Personality Assessment*, *60*(1), 48-59. doi:10.1207/s15327752jpa6001_3

Humphris, G., & Kaney, S. (1998). The encouragement of perfect health professionals. *Medical Education*, *32*, 452-455.

Ioakimidis, M. (2007). Online or on campus: A student tertiary education cost model comparing the two, with a quality proviso. *Higher Education in Europe*, *32*(2&3), 249-260.

Ioakimidis, M. (2010). Home sweet home: Precluding two sources of academic stress in higher education through distance/e-learning. *Georgian Electronic Scientific Journal: Education Science and Psychology*, *2*(17), 29-39.

Ives, S.K. (2010). *The impact of an online orientation program on the imposter phenomenon, self-efficacy, and anxiety* (Doctoral dissertation). Retrieved from ProQuest Digital Dissertations database. (AAT 3433523).

Kolligian, J., & Sternberg, R. (1991). Perceived fraudulence in young adults: Is there an 'impostor syndrome'? *Journal of Personality Assessment*, *56*(2), 308-326. doi:10.1207/s15327752jpa5602_10

Langford, J., & Clance, P. (1993). The imposter phenomenon: Recent research findings regarding dynamics, personality and family patterns and their implications for treatment. *Psychotherapy: Theory, Research, Practice, Training, 30*(3), 495-501. doi: 10.1037/0033-3204.30.3.495

Marmor, J. (1998). Anxiety. *The encyclopedia of mental health*, *1*, 212-216. Franklin Watts. doi: 10.1037/11558-011

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal & Social Psychology*, *67*(4), 371-378.

Milgram, S., Bickman, L., & Berkowitz, L. (1969). Note on the drawing power of crowds on different size. *Journal of Personality and Social Psychology*, *13*, 79-82.

McGregor, L., Gee, D., & Posey, K. (2008). I feel like a fraud and it depresses me: The relation between the imposer phenomenon and depression. *Social Behavior & Personality: An International Journal, 36*(1), 43-48.

Montgomery, R.L. (1992). Social influence and conformity: A transorientational model. In D. Gronberg & G. Sarup (Eds.), *Social Judgment and Intergroup Relations* (pp. 175-200). New York, NY: Springer-Verlag.

Moore, H.T. (1921). The comparative influence of majority and expert opinion. *The American Journal of Psychology*, *32*(1), 16-20. Retrieved from http://www.jstor.org/stable/1413472

Nepon, T., Flett, G.L., Hewitt, P.L., & Molnar, D.S. (2011). Perfectionism, negative social feedback, and interpersonal rumination in depression and social anxiety. *Canadian Journal of Behavioral Science*, *43*(4), 297-308.

Nesdale, D., & Pelyhe, H. (2009). Effects of experimentally induced peer-group rejection and out-group ethnicity on children's anxiety, self-esteem, and in-group and out-group attitudes. *European Journal of Developmental Psychology*, *6*(2), 294-317. doi:10.1080/17405620601112436

Offstein, E.H., Larson, M.B., McNeill, A.L., & Mwale, H.M. (2004). Are we doing enough for today's graduate student? *The International Journal of Educational Management*, *18*(6/7), 396-407.

Parks, M.R., & Floyd, K. (1996). Making friends in cyberspace. *Journal of Communication*, *46*(1), 80-97.

Patterson, B., & McFadden, C. (2009). Attrition in online and campus degree programs. *Online Journal of Distance Learning Administration*, *12*(2). Retrieved from http://www.westga.edu/~distance/ojdla/summer122/patterson112.html

Penn State College of Agricultural Sciences. (2012). How to create a PSU Listserv. Retrieved from http://agsci.psu.edu/it/how-to/how-to-create-a-psu-listserv.

Price, V., Nir, L., & Cappella, J. N. (2006). Normative and Informational Influences in Online Political Discussions. *Communication Theory (10503293)*, *16*(1), 47-74. doi:10.1111/j.1468-2885.2006.00005.x

Reynolds, W. M., Dowell, S. R., Peeters, L. A., Beene, K. E., Espino, A. E., & Longstreth, D. (2007). *Perfectionistic Self-Presentation Scale: Psychometric Characteristics in College Students*. Washington, District of Columbia, US: American Psychological Association (APA).

Rhodewalt, F. (1990). Self-handicappers: Individual differences in the preference for anticipatory, self-protective acts. *Self-handicapping: The paradox that isn't* (pp. 69-106). New York, NY: Plenum Press.

Ross, S.R., Stewart, J., Mugge, M., & Fultz B. (2001). The imposter phenomenon, achievement dispositions, and the five factor model. *Personality and Individual Differences, 31*, 1347-1355.

Sharpley, C.F., Bitsika, V., & Christie, D.R.H. (2010). Why I feel bad: Refinement of the effects of prostate cancer upon lifestyle questionnaire and an initial exploration of its links with anxiety and depression among prostate cancer patients. *Psycho-Oncology*, *19*, 839-846.

Schultz, P.W., Khazian, A.M., & Zaleski, A.C. (2008). Using normative social influence to promote conservation among hotel guests. *Social Influence*, *3*(1), 4-23. doi: 10.1080/15534510701755614

Snyder, T.D., & Dillow, S.A. (2011). Digest of education statistics 2010. U.S. Department of Education. Retrieved from www.nces.ed.gov/pubs2011/2011015.pdf

Sonnak, C. & Towell, T. (2001). The imposter phenomenon in British university students: Relationships between self-esteem, mental health, parental rearing styles and socioeconomic status. *Personality and Individual Differences, 31*, 863-874.

Stein, E.L., & Weidman, J.C. (1989, November). *Socialization in graduate school: A conceptual framework*. Paper presented at the Annual Meeting of the Association for the Study of Higher Education, Atlanta, GA.

Suler, J. (2004). The online disinhibition effect. *CyberPsychology & Behavior*, *7*(3), 321-326.

Thompson, T., Davis, H., & Davidson, J. (1998). Attributional and affective responses of impostors to academic success and failure outcomes. *Personality and Individual Differences, 25*(2), 381-396. doi:10.1016.S0191-8869(98)00065-8

Tolunay, A. (2005). *Group identify effects on social influence* (Doctoral dissertation). Retrieved from ProQuest Digital Dissertations database. (AAT 3186924).

Topping, M., & Kimmel, E. (1985). The imposter phenomenon: Feeling phony [Abstract]. *Academic Psychology Bulletin, 7*(2), 213-226.

Triplett, N. (1898). The dynamogenic factors in pacemaking and competition. *American Journal of Psychology*, *9*, 507-533. Retrieved from http://psychclassics.yorku.ca/Triplett/

U.S. Department of Education, National Center for Education Statistics. (2012). *Digest of Education Statistics, 2011* (NCES 2012-001). Retrieved from http://nces.ed.gov/fastfacts/display.asp?id=98

Weisband, S.P., Schneider, S.K., & Connolly, T. (1995). Computer-mediated communication and social information: Status salience and status differences. *Academy of Management Journal*, *38*(4), 1124-1151.

Wolniewicz, R. (1996, November). *The psychology of the graduate student: A preliminary study*. A paper presented at the Annual Meeting of the Speech Communication Association, San Diego, CA.

Zung, W.W.K. (1971). A rating instrument for anxiety. *Psychosomatics*, *12*(6), 371-379

Information Sharing, Community Development, and Deindividuation in the eLearning Domain

Nicole A. Cooke
University of Illinois at Urbana-Champaign

Abstract
In a study of the information behaviors of graduate students enrolled in an online Masters of Library and Information Science (MLIS) program, it was determined that learners engage in threaded discussions not only for cognitive purposes but for affective reasons as well. The information sharing among students was particularly prolific during a session in which medical ailments and information were discussed. Data were collected from an asynchronous class in a graduate LIS program, and were examined through learner/context analysis and textual analysis. This study used syllabi, course construction, specific assignments and requirements, and other details that contribute to the totality of the learning environment. Specific attention was given to the threaded discussions assigned in the class. This data provided insight into the students' activities and learning during 15 weeks, and enhanced the overall context for the small world that develops within an online learning community.

Students connected with their peers and instructor through copious exchanges of information during which a concerted and consistent effort was made to connect with one another by using personal names, engaging in humor and joke-telling, using emoticons, and expressing support and empathy.

Introduction

In a study of the information behaviors of graduate students enrolled in an online Masters of Library and Information Science (MLIS) program (Cooke, 2012), it was determined that learners engage in threaded discussions not only for cognitive purposes but for affective reasons as well. Students connected with their peers and instructor through copious exchanges of information (information sharing) during which a concerted and consistent effort was made to connect with one another by using personal names, engaging in humor and joke-telling, using emoticons, and expressing support and empathy. Referred to as *connecting* (Cooke, 2014), learners in these online environments initiated and maintained relationships, and developed community over the course of a semester by sharing information with one another. Connecting represents learners' attempts to interact with one another on a personal level by significant use of personal names, expressions of support, empathy, humor, and also evidence of instructor immediacy.

This study addressed the following research questions: What patterns of information interactions are exhibited in the written interactions of the graduate students in an online learning community? What impact, if any, does the context of a small world community have on the information behaviors of online students? This research examined two courses, which are referred to as *Technology* and *User Studies*. User Studies is a theoretical class examining people's information seeking, searching, using, and valuing behaviors, and their impact on services provided by libraries and information organizations. Technology is a practical, hands-on class that introduces students to key concepts about the Internet, programming, and selected hardware and software that future library professionals may encounter, and examines their role in library services. In this graduate LIS program, classes are 15 weeks in length, and students typically take the two classes selected for this research concurrently in their first semester of study.

Literature Review

Information Behavior

Small worlds. The formation of community within the particular setting of online learning is attributed to the concepts of "small worlds," an information behavior theory posited by library and information science (LIS) scholar Elfreda Chatman (1991). Small worlds encompass several concepts including normative behavior; worldview; social types; information behaviors; and social network theory (Case, 2012; Yu, 2012; Huotari & Chatman, 2001; Chatman, 2000; Chatman, 1999; Chatman, 1991). Huotari and Chatman described the theory in the following way:

> In addressing the theory of small worlds, it is essential to remember that the reason the small world works is that it allows persons to share a similar cultural and intellectual space. That is, those things that hold this world together include a common assessment of information worthy of attention, social norms that allow its members to approach or ignore information and behaviors that are deemed by other inhabitants to be appropriate for this world. (Huotari & Chatman, 2001, p. 352)

Students enrolled in online MLIS programs share the culture established by the overall library community, and they share an intellectual space and corresponding academic culture by engaging in a formal learning environment. Furthermore, by studying a specific subject area, in this case LIS, students seek and share specific information that assumes and promotes social norms, normative behaviors, and worldviews appropriate to the field of librarianship (Burnett & Jaeger, 2011; Burnett et al., 2001; Chatman, 2000; Chatman, 1999). Online communities, especially those revolving around academic content, provide a common set of interests for their members.

Affective Information Sharing

Connecting is related to the affective or emotional realm of information behavior. Literature about the affective domain of information behavior is growing, and includes the feelings and mental states of users as they seek, use, and avoid information (Cibangu, 2015; González-Ibáñez, 2015; Savolainen, 2015a; Savolainen, 2015b; Fourie & Julien, 2014; Nahl & Bilal, 2007). Of particular note are the foundational works of the following information science scholars. Nahl's (2005, 2004, 2001) work addresses users' feelings of frustration, impatience, information overload, resistance to new information, and confusion. Mellon (2015, 1986) discusses feeling of anxiety, while Harris, Stickney, Grasley, Hutchinson, Greaves, and Boyd (2001) address disappointment in relation to information seeking. Kuhlthau (1993) investigates the feeling of uncertainty, often expressed as anxiety or worry, and Heinström (2014, 2004) also discusses stress, worry, and feelings of low confidence in information consumers. The findings of this research and their relation to this literature provides a natural link back to Chatman's (1992; 1996) work that address the small worlds of insiders and outsiders and retired women, and the emotions and feelings resulting from their information seeking and use.

Community Development

Community was developed both inside and outside the formal course environment. This coalescence of connections enabled students to work with the course content collectively, and this group engagement made each class a community of practice (albeit one with a finite life cycle).

> It is now generally accepted that people engaging in electronic exchanges are able to create communities—places with socially constituted norms, values, and expectations. Text serves as the lifeblood of these electronic places, conveying the ideas and feelings of participants that lead to the growth and evolution of a community or to its demise. … A virtual community is comprised of members 'bound together for mutual service'. Members of virtual communities tend to provide advice and solutions to problems expressed by other members, even though they may be strangers to one another. (Burnett et al., 2003, paras. 1 & 5)

As some students expressed, forging connections and creating community in an online environment can be difficult to accomplish due to the anonymity, asynchronicity, and lack of personal interaction and visual cues. Students said it was difficult to communicate solely by text, and expressed the desire to see their classmates. Forming relationships and a small world in this environment is different and perhaps more challenging than doing so in a face-to-face environment (Haythornthwaite & Kendall, 2010; Kazmer, 2007, 2005; Haythornthwaite et al., 2000; Baym, 1997, 1995) and requires effort, risk-taking, and a willingness to trust (Finlay & Willoughby, 2008). Interacting in the threaded discussions enabled the students to engage with course content while interacting with one another. The resulting conversations allowed students to "… publish, reflect, discuss, critique, and connect their knowledge" (Finlay & Willoughby, 2008, p. 54).

Distance Education

The distance education literature (particularly within information science) addresses community development among online learners, especially as developed through computer-mediated communication (Trespalacious & Rand, 2015; Poole 2013; Yukawa, 2010; Haythornthwaite & Hagar, 2004; Haythornthwaite et al., 2000; Haythornthwaite, 2001; Kazmer, 2010, 2006, 2000). Since online learners do not have physical access to their instructors and fellow students, they must be purposeful in their interactions and efforts to make contact with one another. Socializing becomes a function facilitated by technology. Kazmer (2000) stated that forming community is an important coping skill for distance students.

> They are in a new and unfamiliar learning environment, without physical classroom and with limited face-to-face contact. They face a variety of problems, social and technological, that students in more traditional programs do not. As students enter this new learning environment, they need support to help them gain entry to the community and to begin their interaction with others. (Kazmer, 2000, p. 2)

Haythornthwaite (2000) suggested that not only is community building important for distance learners, but community maintenance is vital as well. Technology facilitates community building, but students must make a concerted effort to maintain and nurture the initial bonds formed. Disengaging and not maintaining the social bonds and connections is referred to as "fading back" (p. 12). "Those who fail to make such connections feel isolated and more stressed than those who are more active in the community" (p. 2).

Palloff and Pratt (2001) state: "It is always important to remember that in the online environment, we present ourselves in text. Because it is a flat medium, we need to make an extra effort to humanize the environment" (p. 18). In the online classroom students interact exclusively via text, e-mails, journals, assignments, and threaded discussions. In this environment, it is important to promote social presence

(Song & Yuan, 2015; DeSchryver et al., 2009; Liu et al., 2009; Scollins-Mantha, 2008; Kehrwald, 2008; Ouzts, 2006; Biocca et al., 2003; Rettie, 2003; Richardson & Swan, 2003; Stein & Wanstreet, 2003; Tu, 2000; Tu & McIsaac, 2002; Gunawardena & Zittle, 1997). Social presence is defined as "… the degree of salience of the other person in the (mediated) interaction and the consequent salience of the interpersonal relationship. This is interpreted as the degree to which a person is perceived as 'real' in mediated communication" (Richardson & Swan, 2003, p. 70). This realness can be thought of as "… the degree to which a user feels access to the intelligence, intentions, and sensory impressions" of the other members of the online environment (Tu, 2000, p. 28). Social presence needs to be cultivated, varies from group to group, depends on the particular technologies available to the learners, and the culture of the group in question (Gunawardena & Zittle, 1997). Social presence is an important element in an online learning environment because of the lack of nonverbal and other interpersonal cues that are fundamental to face-to-face interactions in classrooms. Online cues and interactions are strategies learners use to overcome transactional distance, to get to know one another, and to form the basis for community in the online environment.

Methods and Analysis

Approaches to the Research

Informed by the constructs of phenomenography and virtual ethnography, this research employed textual analysis to examine the data collected from two online classes. The goals of this research were, in part, to: 1) Examine the dynamics of information behaviors in an asynchronous online classroom and, 2) identify factors that shape these behaviors. The study addressed the question: What information behavior patterns, if any, do students in an online asynchronous learning communities exhibit?

Two courses were selected and the respective instructors were asked for assistance with this study. Both instructors allowed the researcher access to their online course shells after the semester had ended. Each class contained 19 students. No incentives for participation were offered to the instructors or students. Data were collected from two online and asynchronous courses, *Technology* and *User Studies*, at an ALA accredited program in Library and Information Science at a university in the northeast United States. Data for this study consisted of the work product from these two classes (threaded discussions) and were collected over the course of one full academic semester (15 weeks) and examined through a textual analysis. The researcher did not participate in either class, but instead became immersed in the online course shells after the courses were completed.

This study was qualitative in nature, and was informed by phenomenography, virtual ethnography, and naturalistic inquiry. In order to answer the research question, it was necessary to become immersed in the small worlds of the LIS graduate students and investigate their experiences within the online learning environment. These approaches to qualitative research facilitated the discovery of students' information behaviors and the meanings associated with those behaviors. A phenomenographic approach is beneficial for uncovering and unpacking community members' experiences and analyzing the information exchanges and community development that occurred within the small world environment. Phenomenography was employed in this research as an interpretive method of identifying and unpacking the experiences of graduate students in the User Studies and Technology classes. The text-based discussions revealed learners' feelings of achievement, frustration, and community as they progressed through the semester.

Similar to traditional ethnography, virtual ethnography requires the researcher to be a member of and participant in the cyber culture (Rybas & Gajjala, 2007; Teli et al., 2007) or online community being investigating (Hine, 2000, 2008; Paccagnella, 1997; Ward, 1999). Virtual ethnography lends itself to the immersive and extended study of online learning communities (Rutter & Smith, 2005) and small worlds (Chatman, 1991).

This research used ethnographic/virtual ethnographic techniques. The main techniques are naturalistic inquiry and thick description of the participants' online learning environment, following the development of the participants' social interactions over time (Joinson, 2005) and identifying participants' patterns of information behavior through post-course immersion in the data. The researcher accordingly became immersed in the online course shells as an observer, not as a community participant, and observed the natural occurrences that transpired during a semester-long online course (Rutter & Smith, 2005). The researcher also investigated "… the complex interaction between trust, intimacy, disclosure and time as complex relationships develop" (Carter, 2005, p. 149), and the influence of these relationships on the development of community and information exchange in the online classroom.

Phenomenography and virtual ethnography have qualities in common such as being immersive, highlighting members' meanings and community, and being context- specific. The combination of these two approaches enabled a specific lens through which to study the information behavior of learners in a small world. Phenomenography, an approach used in face-to-face study, unites with virtual ethnography to examine the specific needs and characteristics of an online learning environment. The two merge to create an atmosphere conducive to naturalistic inquiry and elucidated a set of principles that guided the researcher. The researcher specifically engaged in immersion and close observation of particapants and their lived experiences by examining the totality of the students' online activites as represented by threaded discussions.

In order to approach this research from a naturalistic perspective (Crystal & Wildemuth, 2009; Erlandson, Harris, Skipper & Allen, 1993; Lincoln & Guba, 1986; Miles & Huberman, 1994), and collect and analyze data that "… more closely reflect the real, lived experiences of the population of interest" (Crystal & Wildemuth, 2009, p. 62), textual analysis was used for data analysis. This naturalistic approach is more context- specific than content analysis, and these methods enabled the researcher to elucidate new areas of human information behavior.

Textual Analysis

Threaded discussions are the most important and plentiful components of an online class as "… discussions can create a mutual sense of interaction and belonging that is essential to feeling the social presence of others" (Rovai, 2007, p. 103). Large volumes of content are produced in a short time, and because multiple students generate content, evaluating threaded discussions is challenging. "Evaluating online discussions is neither as simple nor as straightforward as one might suppose; it involves answering important questions about the instructor's purpose, the student learning to be measured, and the application of coding procedure" (Meyer, 2006, p. 83).

Discussion threads are the "media" through which the information behavior of online learners was ascertained (Fairclough, 2003, p. 30), and documents and forms of material culture examined (Lindlof & Taylor, 2010). Lindlof and Taylor characterize material documents as "mute evidence" that cannot respond to researcher questioning, yet are rich sources of information that can be used to understand participants and phenomena of interest. Within material documents are critical incidents (Flanagan, 1954) that explicate the information behaviors and intents of participants.

> Text, objects, and spaces do have a lot to 'say' when we read them alongside the living voices of informants and other social actors. Moreover, people do disclose their understandings of, and feelings about, the material world in other ways besides introspection – for example, by gesture, posture, facial expression, stories and accounts, jokes, ironic asides, confessions, even silence, (Lindlof & Taylor, 2010, p. 271)

Many forms of material culture are found in the threaded discussions, and to analyze this significant source of information, textual analysis (Krippendorff, 2004; McKee, 2003; Neuendorf, 2002; Spurgin & Wildemuth, 2009) was employed. Similar to what was described by Lindlof and Taylor, McKee defined textual analysis as:

> A way for researchers to gather information about how other human beings make sense of the world. It is a method … for those researchers who want to understand the ways in which members of various cultures and subcultures make sense of how they are, and of how they fit into the world in which they live. (2003, p. 1)

Textual analysis is an interpretive approach that facilitated the discovery of information interactions, intents, flow, learning, connecting and the development of community, as they emerged in students' discussions. Texts, in the form of threaded discussions, provided insight into the learners' experiences and the meanings assigned to them. McKee emphasized the benefits of textual analysis by suggesting that "… the reason we analyze texts is to find out what were and what are the reasonable sense-making practices of cultures, rather than just repeating our own interpretation and calling it reality" (p. 19).

Theory for Textual Analysis

Textual analysis was conducted through the lens of the information behavior theory *information intents* (Todd, 2005). Information intents suggests that people seek and acquire knowledge to *Get a Complete Picture,* to *Get a Changed Picture,* to *Get a Clearer Picture,* to *Get a Verified Picture,* and to *Get a Position in the Picture* (Todd, 2005, pp. 198-203). Newly acquired information adds to an individual's existing knowledge base and facilitates the expansion of viewpoint. Information intents allows information behavior patterns to be discovered.

To achieve consistency in the coding scheme, coding was conducted in three rounds, or exposures, occurring over a 60-day period. A total of 33 discussion threads were analyzed, 30 from the User Studies class, and three from the Technology class. (The Technology class only required three graded discussions.) All threads were analyzed with NVivo software, a tool for qualitative data analysis (http://www.qsrinternational.com). Codes used to analyze the threaded discussions were derived directly from the information intents theory (Todd, 2005) and from the results of a previous study that yielded a code named ***Getting Connected*** (Cooke, 2014). The codes were assigned to text in the threaded discussions and counted to determine the numbers of statements made over the course of the semester. Each thread was examined line-by-line, and codes were assigned to portions of the text in the discussions, ranging from a few words to a few sentences. A constant comparative approach was uses, always comparing new information with previously identified information, identifying patterns, refining coding assignments as needed, and developing new coding categories as needed (Glaser & Strauss, 1967).

Results

The major component/source of data for this online class was the threaded discussions. Discussions were compulsory and comprised a portion of students' overall course grades. Students were given guidelines for participating in the discussions (one original post in each thread and two responses to classmates' postings, all of which should be substantive and draw from content contained in the lectures and course readings). Since students' writing styles vary, the length of entries varied. Some primarily used bullet points, while others wrote their posts as though the assignments were academic essays.

The threaded discussions were the bonding text, or "glue" in the classes and were the place of interactions between students and instructor. The students came together through discussions that incorporated course content, lectures provided by the professor, and scholarly articles about their topics of study. The weekly discussion thread questions were straightforward and asked students to compare and contrast the theories presented in their weekly readings, as well as discuss the implications of the theories for their role as library and information professionals.

A significant result of the research is the aforementioned element of connecting (Cooke, 2014). Many examples of connecting were evident in the threaded discussions of the two graduate classes. Written demonstrations of the online learners connecting with one another included:

- *I agree, XXX.*
- *Like XXX, I think YYY has hit the nail on the head in her post.*
- *This is an excellent point, XXX.*
- *XXX, you're so funny! :)*
- *Thank you, XXX, for the link to the article that defines the different terms related to "construct."*
- *XXX, your library rocks!*
- *Wow, XXX - out of curiosity, how does this work logistically? Is somebody actually up all night answering queries? There's a budget for that?*

Another dimension of *Connecting* was the affinity students felt for the instructor in the User Studies course, affinity that aided in the formation of community and facilitated understanding of course content. The instructor was a consistent source of comfort, encouragement, and humor, while providing more traditional functions such as giving feedback about performance, answering questions and providing correction as needed. The instructor maintained a strong presence throughout the semester in the threaded discussions.

Health Information Sharing

A specific and notable example of intense information sharing in the study involved the participants' discussion of health and medical information. Even though students were expressly told by the instructor not to disclose personal medical conditions or experiences in the threaded discussions, students shared an inordinate amount of personal information about themselves and family members, revealing a wide variety of personal conditions. Students revealed their ages and personal conditions including: stage four lung cancer, scleroderma, torn ACL ligaments, pregnancy and bed rest, pulmonary fibrosis, sinus conditions, spinal injuries, heart conditions, appendicitis, acute myelogenous leukemia, fibromyalgia, Lyme disease, drug abuse prevention measures, multiple sclerosis, osteochondritis, caring for elderly parents, and assorted gynecological issues.

Related discussion thread posts include:

- *People who lived a life in the round were often seen to be information poor because their information came from a very small group of people. When thinking about it in that way, the fact that weak ties provide more and/or different information makes sense. People with whom we have weak ties live in a different circle and have a different knowledge base, like the nurses in Pettigrew's article. When we're sick we may be very tempted to only talk to our closest friends and family, but it may actually be more beneficial to branch out to others.*
- *Recently I was diagnosed with a very common problem that luckily does not include any scary outcomes. Even though my doctor reassured me that I needn't worry I couldn't really concentrate on what he was saying after he told me what was going on. So even though it was good to speak with him, and eventually I relaxed enough to hear what he*

was saying, it would have been nice if he could have directed me to some written information that I could take home with me, or even a trusted website that I could read at a later time. This way I could have gotten some more in-depth information at a time when I was better able to comprehend it.

- *How can health care professionals use this knowledge to create similar atmospheres in their own practices? Perhaps waiting rooms could be turned into information grounds, not just by having pamphlets lying around that no one reads, but with nurses who are available to answer questions before and after a visit to see the doctor?*
- *Seven years ago, my back surgeon offered me a list of contacts who had previously undergone spinal fusion. The phone call provided a comfort and someone I could relate to. I felt like I was in a special club because no one in my circle of family/friends could share my experience. I have never recognized the value of this support group nor how my doctor valued his patients.*
- *So sometimes it's not just medical knowledge one needs, it's help navigating the system.*
- *We truly have to be our own patient advocates and cannot simply rely on information we get from the Internet.*

In addition to having relevant topics to which participants can relate, common experiences proved to be a significant discussion generator. Everyone has had some type of experience with illness, and the need for quality and accessible medical information is universal. The relative anonymity of the asynchronous environment likely provided a layer of security and comfort that made students willing to share such personal information. Though students knew each other to some degree by the time this discussion occurred in the semester, their ignorance of each other's appearance combined with the fact that they did not have to see each other as they "spoke" probably made sharing personal medical details a non-threatening experience. In fact, this environment may have served to build an even stronger sense of community among this group and thereby permitted a deeper level of sharing.

Johnson and Case (2012) discuss the demographic characteristics of people who seek health information. Of particular interest is the characteristic of personal experience (pp. 51-56). The relevance or salience of the health information and the seeker's "proximity" (p. 55) to the information can influence the likelihood of sharing said found information with others. Information sharing can be a type of coping mechanism and a means of garnering social support (pp. 69-74), but health information is a common denominator with people, as everyone has a health issue in their immediate sphere and can relate to others in similar situations. This universality can hasten and amplify information sharing and community development.

These findings relate directly to this study's proposed research question that sought to determine the patterns exhibited in the written interactions of graduate students in an online learning community, and links this question to the idea that community formation influences the information behavior and sharing of students. Patterns of affective information exchange indicate the formation of community among students and suggest that the affective dimension of information sharing contributes to increased levels of information exchange and conversation in the discussions. Findings also indicated the interconnection of the cognitive and affective dimensions of the information intents schema. Information behavior and sharing are not compartmentalized into cognitive and affective dimensions, rather these dimensions complement one another to form a holistic and comprehensive view of information seeking, utilization, and sharing, and they illustrate how emotion is represented in text and contributes to the overall production of knowledge.

Discussion

Findings from this study relate to the bodies of work treating the affective dimension of information behavior, distance education, and social psychology. Findings indicate that the phenomena of interest in this study also have roots in other areas of the literature, some outside the field of information behavior and library and information science. Drawing upon literature in psychology yields insight into the motivations that possibly undergird the type of information sharing that occurred around medical / health information. The medical / health information sharing that occurred in this study could be considered over-sharing, yet it seemed to facilitate some of the community development that occurred in the online learning environment. Specifically, the literatures on social sharing, help seeking, and disinhibition and deindividuation contribute to the understanding of these especially personal types of information sharing.

Social Sharing

Sharing is socially constructed, particularly when shared within digital and social networks, and can be fraught with politics, emotion, and other cultural elements—simply stated, sharing is "complicated" (Wittel, 2011, p.5). As an example of the complexities of sharing, John (2013) describes an aspect of information sharing that involves the sharing of personal details and feelings. In the class' threaded discussions, students primarily discussed academic texts and tasks, and their aspirations for their impending professional lives. The discussion threads focused on medical and health information behavior allowed them an opportunity to share personal details and feelings. Arguably, this was the first opportunity to be social with one another beyond a cursory level, particularly given the asynchronous nature of the class. In addition to revealing actual diagnoses, the discussion participants shared a substantial amount of information about their *feelings* about said diagnoses and ailments. For example, one student described being overwhelmed and not being able to relax to absorb the doctor's advice. Other students discussed welcoming and trusted medical professionals and environments. Student comments indicate that the contextual environment can be just as important as the information being received. These scenarios were common to the people in the class and established commonality that facilitated discussion and understanding.

In this way, information sharing is an act of communication distribution. Communication and information exchange are critical components of community building. John (year?) suggests that this kind of sharing is a result of people having "something in common" (p. 114) (in this case, medical conditions) and is a way of "imparting one's inner state to others" (p. 115). This social aspect of sharing, the sharing of feelings and emotion, can also be considered an important part of community building; affective sharing implies that a level of respect and trust exists within a group, and also serves to strengthen those bonds of trust and community.

Help Seeking

Help seeking appears in several bodies of literature, including information science, nursing, public health, and psychology (Abrahamson et al., 2008; Galdas, et al., 2005; Nicholas et al., 2004; Snowden, 1998; Ybarra & Suman, 2006), and generally speaking, describes the phenomenon of an individual seeking medical / health information from the Internet or from a trusted source. Trusted sources can include both formal sources (medical professionals or counselors) or informal (peers or family), and the information being sought could be for the individual or for someone being cared for by that individual (known as Lay Information Mediary Behavior) (Abrahamson et al., 2008).

The literature suggests that help seeking is a secondary level of information solicitation and is used to confirm an existing diagnosis or self-diagnosis. Help seeking is also a means of seeking social support; social support eases pain and isolation for both those experiencing the ailment and for those caring for them. Ybarra and Suman (2006) suggest that this is the most important part of help seeking, more important than the validation aspect. They state "social support is a necessary component of somatic and mental health" (p. 38). The social support that results from help seeking can lessen stress and worry, can improve cognitive understanding of the medical issue at hand, can even lead to lifestyle changes (e.g., smoking cessation) (Abrahamson et al., 2008, p. 317), and can increase the resilience of both patients and caregivers (Nicholas et al., 2004, p. 17). One student in the study mentioned reaching out to a doctor and receiving a comforting phone call and additional information resources about their ailment, and indicated how valuable that kind of support was on a personal level.

Help seeking for social support could also provide an explanation for the high level of personal information sharing that occurred in the discussion threads. Learners shared personal medical / health information not only because it served as a point of commonality with their peers, but because it was an opportunity to validate their own diagnoses and situations and even gain new information about resources, cures, treatments, etc. Ultimately, this type of information sharing was perhaps solely an attempt to lean on a virtual shoulder of support and lessen their own stress and anxiety.

Deindividuation

The social psychology literature suggests that this type of over-sharing of information, a break from normative behavior, demonstrates a lack of inhibition called deindividuation (Suler, 2004; Coleman et. al, 1999; Zimbardo, 1970). This lack of inhibition or self-awareness in a specific setting can be attributed to the perception that individual group members are "not seen or paid attention to" (Festinger et al., 1952, p. 382), rather personal information was being disclosed to the group-at-large. Though students knew each other to some degree because of ongoing information sharing during the semester, their ignorance of each other's appearance combined with the fact that they did not have to see each other as they "spoke" likely made sharing personal medical details a non-threatening experience. In fact, this environment may have served to build an even stronger sense of community among this group and thereby permitted a deeper level of sharing. This sharing also enabled students to contextualize their own learning and facilitated the construction of new knowledge.

Patterns of help-seeking, deindividuation, social sharing, in conjunction with affective information behavior, indicate the formation of community among students, and suggest that this type of information sharing contributes to increased levels of interaction and conversation. Information behavior is not compartmentalized into cognitive and affective dimensions; rather, these dimensions complement one another to form a holistic and comprehensive view of information sharing and use, and illustrate how emotion contributes to the overall learning process.

Conclusion

The study looked at 38 online graduate students and this population yielded considerable data through the threaded discussion. Despite the rich results from the study there were several limitations. The study was limited by sample size, limited duration of data collection, concentration on one mode of online learning (asynchronous), and the examination of one LIS master's program. While interesting results about health information sharing emerged, that specific phenomenon was not the focus of the study; future research should focus on this phenomenon explicitly and should delve deeper into the possible theories behind medical / health information sharing. There is a wealth of literature about the health information seeking that happens on the Internet, but less of that literature focuses on information sharing and the motivations behind that sharing, especially in enclosed online environments (i.e., the online learning space). Future study should also examine face-to-face classrooms to see if comparable medical / health information sharing occurs in that setting.

This research provided a basis for understanding how students in online LIS courses create connections and build community over the span of a semester. Specifically, an example of health information sharing within threaded discussions highlighted the small world development that can occur in an online learning environment. This examination, derived from Chatman's theory of small worlds (1991), included insights from studies in information science, education, psychology, and communication, and shed new light on online learners—their information behaviors and patterns of information sharing. Further, the instructional design of LIS distance education can benefit from these results and lead to more productive pedagogies in online classes. The insights gained from this research benefit not only the discipline of LIS, but also all others that utilize distance education technologies. Learning online is not just about delivering course content in an online format, but rather a way to put the learner first in the course design process by considering how they learn best and facilitating the development of learning communities.

Ultimately, small worlds do influence the information behavior of online students. Small worlds are forged through interaction and exchanges of information. These interactions increased the sense of community felt by students, and in turn this sense of community encouraged more interaction. In this way, small worlds are cyclical and dynamic entities. Small worlds are developed around context and depend upon interaction, norms, cognitive and affective information seeking and sharing, and can foster deep and sustained learning and construction of knowledge.

As indicated by the variety of opinions about information sharing, in multiple disciplines, sharing as it pertains to online learning and community development is a complicated phenomenon. This study begins to address the questions of what information sharing patterns students in online asynchronous learning communities exhibited and what role the formation of community played in the online learning environment. Learners in this online environment initiated and maintained relationships, and developed community over the course of a semester by purposely and freely sharing information with one another.

References

Abrahamson, J. A., Fisher, K. E., Turner, A. G., Durrance, J. C., & Turner, T. C. (2008). Lay information mediary behavior uncovered: Exploring how nonprofessionals seek health information for themselves and others online. *Journal of the Medical Library Association: JMLA*, *96*(4), 310-323. DOI: 10.3163/1536-5050.96.4.006

Baym, N. K. (1995). The emergence of community in computer-mediated communication. In Steve Jones (Ed.) *CyberSociety: Computer-mediated communication and community* (pp. 138–63). Thousand Oaks: Sage.

Baym, N. K. (1997). Interpreting soap operas and creating community: Inside an electronic fan culture. In Kiesler, S. (Ed.) *Culture of the Internet* (pp. 103–20). Mahwah , NJ: Lawrence Erlbaum.

Biocca, F., Harms, C., & Burgoon, J. K. (2003). Criteria for a theory and measure of social presence. *Presence: Teleoperators & Virtual Environments, 12*(5), 456–480. DOI: 10.1162/105474603322761270

Burnett, G., Dickey, M.H., Kazmer, M.M., & Chudoba, K.M. (2003). Inscription and interpretation of text: A cultural hermeneutic examination of virtual community. *Information Research*, *9*(4) paper 162. Retrieved from http://www.informationr.net/ir/9-1/paper162.html

Burnett, G., & Jaeger, P. T. (2011). The theory of information worlds and information behaviour. *New Directions in Information Behaviour*, *1*, 161.

Carter, D. (2005). Living in virtual communities: An ethnography of human relationships in cyberspace. *Information, Communication & Society, 8*(2), 148-167.

Case, D. O. (2012). *Looking for information: A survey of research on information seeking, needs and behavior*. Emerald Group Publishing.

Chatman, E. A. (2000). Framing social life in theory and research. *New Review of Information Behaviour Research, 1*, 3-17.

Chatman, E. A. (1999). A theory of life in the round. *Journal of the American Society for Information Science, 50*(3), 207-217.

Chatman, E.A. (1996). The impoverished life-world of outsiders. *Journal of the American Society for Information Science, 47*(3), 193-206.

Chatman, E.A. (1992). *The information world of retired women*. Westport, CT: Greenwood.

Chatman, E. A. (1991). Life in a small world: Applicability of gratification theory to information-seeking behavior. *Journal of the American Society for Information Science, 42*(6), 438-449. DOI: 10.1002/(SICI)1097-4571(199107)42:6<438::AID-ASI6>3.0.CO;2-B

Cibangu, S. K. (2015). A new direction in information science research: making information science a human science. *Information Research*, *20*(3).

Coleman, L. H., Paternite, C. E., & Sherman, R. C. (1999). A reexamination of deindividuation in synchronous computer-mediated communication. *Computers in Human Behavior*, *15*(1), 51-65. DOI: 10.1016/S0747-5632(98)00032-6

Cooke, N. A. (2014). Connecting: Adding an affective domain to the information intents theory. *Library & Information Science Research*, *36*(3), 185-191.

Cooke, N. A. (2012). *The information use behaviors of graduate students in an online learning community* (Doctoral dissertation, Rutgers University-Graduate School-New Brunswick).

Crystal, A. J. & Wildemuth, B. M. (2009). Naturalistic inquiry. In B. M. Wildemuth (Ed.), *Applications of social research methods to questions in information and library science* (pp. 62-72). Westport, CT: Libraries Unlimited.

DeSchryver, M., Mishra, P., Koehleer, M., & Francis, A. (2009, March). Moodle vs. Facebook: Does using Facebook for discussions in an online course enhance perceived social presence and student interaction? *Society for Information Technology & Teacher Education International,* 1, 329-336.

Diener, E. (1977). Deindividuation: Causes and consequences. *Social Behavior and Personality: An International Journal*, *5*(1), 143-155. DOI: http://dx.doi.org/10.2224/sbp.1977.5.1.143

Erlandson, D.A., Harris, E.L., Skipper, B.L., & Allen, S.D. (1993). *Doing naturalistic inquiry: A guide to methods.* Newbury Park, CA: Sage.

Festinger, L., Pepitone, A., & Newcomb, T. (1952). Some consequences of deindividuation in a group. *Journal of Abnormal and Social Psychology, 47,* 382-389.

Fairclough, N. (2003). *Analysing discourse: Textual analysis for social research*. London, UK: Routledge.

Finlay, J. & Willoughby, L. (2008). Exploring Online Learning Relationships: A Case Study in Higher Education. In S. Holland (Ed.), *Remote relationships in a small world* (pp. 53-73). New York: Peter Lang.

Fourie, I., & Julien, H. (2014). IRS, information services and LIS research–a reminder about affect and the affective paradigm… and a question. *Library Hi Tech*, *32*(1), 190-201.

Galdas, P. M., Cheater, F., & Marshall, P. (2005). Men and health help- seeking behaviour: Literature review. *Journal of advanced nursing*, *49*(6), 616-623. DOI: 10.1111/j.1365-2648.2004.03331.x

González-Ibáñez, R. (2015). Affective Dimension in Collaborative Information Seeking. In *Collaborative Information Seeking* (pp. 193-208). Springer International Publishing.

Gunawardena, C.N. et al. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research, 17*(4), 397-431.DOI: 10.2190/7MQV-X9UJ-C7Q3-NRAG

Harris, R., Stickney, J., Grasley C., Hutchinson, G., Greaves, L., & Boyd, T. (2001). Searching for help and information: Abused women speak out. *Library and Information Science Research, 23*(2), 123-141. doi:10.1016/S0740-8188(01)00066-4

Haythornthwaite, C. (2001). Exploring multiplexity: Social network structures in a computer-supported distance learning class. *The Information Society, 17*(3), 211-226. DOI:10.1080/01972240152493065

Haythornthwaite, C. & Hagar, C. (2005). The social worlds of the Web. *Annual Review of Information Science and Technology, 39*, 311-346. DOI: 10.1002/aris.1440390115

Haythornthwaite, C., Kazmer, M.M., Robins, J. & Shoemaker, S. (2000). Community development among distance learners: Temporal and technological dimensions. *Journal of Computer-Mediated Communication, 6*(1). Retrieved from: http://jcmc.indiana.edu/vol6/issue1/haythornthwaite.html. DOI: 10.1111/j.1083-6101.2000.tb00114.x

Haythornthwaite, C., & Kendall, L. (2010). Internet and community. *American Behavioral Scientist, 53*(8), 1083-1094.

Heinström, J. (2014). The emotional valence of information experience: relation to personality and approach to studying. In C. Bruce, K. Davis, H. Hughes, H. Cartridge, & I. Stoodley (Eds.), *Information Experience: Approaches to Theory and Practice* (pp. 275-293). Emerald Group Publishing Limited.

Heinström, J. (2004). Fast surfers, broad scanners and deep divers: Fast surfing, broad scanning and deep diving. *Journal of Documentation, 61(2), 228-247. DOI:* http://dx.doi.org/10.1108/00220410510585205

Hine, C. (2000). *Virtual Ethnography*. London, Sage.

Hine, C. (2008). Virtual ethnography: Modes, varieties, affordances. In N. Fielding, R. M. Lee and G. Blank (Eds.), *The SAGE handbook of online research methods* (pp. 257-268). Los Angeles, CA: Sage.

Huotari, M., & Chatman, E. (2001). Using everyday life information seeking to explain organizational behavior. *Library & Information Science Research, 23*(4), 351-366. doi:10.1016/S0740-8188(01)00093-7

John, N. A. (2013). The Social Logics of Sharing. *The Communication Review, 16*(3), 113-131. DOI: 10.1080/10714421.2013.807119

Johnson, J.D. & Case, D.O. (2012). *Health information seeking*. New York: Peter Lang Publishing, Inc.

Joinson, A. N. (2005). Internet behavior and the design of virtual methods. *Virtual methods: Issues in social reseach on the internet.* In C. Hine (Ed.), New York, NY: Berg.

Kazmer, M. M. (2010). Disengaging from a distributed research project: Refining a model of group departures. *Journal of the American Society for Information Science and Technology*, *61*(4), 758-771.

Kazmer, M. M. (2007). How do student experiences differ in online LIS programs with and without a residency?. *The Library Quarterly*, *77*(4), 359-383.

Kazmer, M. M. (2006, Summer). Creation and loss of sociotechnical capital among information professionals educated online. *Library & Information Science Research*, *28*(2), 172-191. doi:10.1016/j.lisr.2006.03.002

Kazmer, M. M. (2005). Community‐ Embedded Learning. *The Library Quarterly*, *75*(2), 190-212.

Kazmer, M. M. (2000). Coping in a distance environment: Sitcoms, chocolate cake, and dinner with a friend. *First Monday, 5*(9). Electronic journal, no pagination. Retrieved from http://http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/791/700. DOI: http://dx.doi.org/10.5210/fm.v5i9.791

Kehrwald, B. (2008). Understanding social presence in text-based online learning environments. *Distance Education, 29*(1), 89-106. DOI: 10.1080/01587910802004860

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology.* Thousand Oaks, CA: Sage Publications, Inc.

Kuhlthau, C. C. (1993). *Seeking meaning: A Process Approach to Library and Information Services.* Norwood, NJ: Ablex.

Lincoln, Y.S. & Guba, E.G. (1986). But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. *New Directions for Program Evaluation, 30*, 73-84.

Lindlof, T. R. & Taylor, B. C. (2010). *Qualitative Communication Research Methods.* Thousand Oaks, CA: Sage Publications, Inc.

Liu, S. Y., Gomez, J., & Yen, C. J. (2009). Community college online course retention and final grade: Predictability of social presence. *Journal of Interactive Online Learning*, *8*(2), 165-182.

McKee, A. (2003). *Textual Analysis: A Beginner's Guide*. London, UK: Sage.

Mellon, C. A. (2015). Library anxiety: A grounded theory and its development. *College & Research Libraries*, *76*(3), 276-282.

Meyer, K. A. (2006). The method (and madness) of evaluating online discussions. *Journal of Asynchronous Learning Networks, 10*(4), 83-97.

Miles, M. & Huberman, M. A. (1994). Qualitative *Data Analysis: An Expanded Sourcebook.* 2nd Ed. Thousand Oaks, CA: Sage.

Nahl, D., & Bilal, D. (2007). *Information and emotion: The emergent affective paradigm in information behavior research and theory*. Information Today, Inc.

Nahl, D. (2005). Affective and cognitive information behavior: Interaction effects in Internet use. *Proceedings of the American Society for Information Science and Technology*, *42*(1).

Nahl, D. (2004). Measuring the affective information environment of Web searchers. *Proceedings of the 67th Annual Meeting of the American Society of Information Science & Technology, 41*, 191-197. DOI: 10.1002/meet.1450410122

Nahl, D. (2001). A conceptual framework for explaining information behavior. *Studies in Media and Information Literacy Education*, ***1***(2), 1-16. DOI: 10.3138/sim.1.2.001

Neuendorf, K. A. (2002). *The content analysis guidebook.* Thousand Oaks, CA: Sage.

Nicholas, J., Oliver, K., Lee, K., & O'Brien, M. (2004). Help-seeking behaviour and the Internet: An investigation among Australian adolescents. *Advances in Mental Health*, *3*(1), 16-23. DOI: 10.5172/jamh.3.1.16

Ouzts, K. (2006). Sense of community in online courses. *The Quarterly Review of Distance Education*, *7*(3), 285-296.

Palloff, R. M., and Pratt, K. (2001) *Lessons from the Cyberspace Classroom: The Realities of Online Teaching.* San Francisco, CA: Jossey-Bass.

Paccagnella, L. (1997). Getting the seats of your pants dirty: Strategies for ethnographic research on virtual communities. *Journal of Computer‐Mediated Communication, 3*(1). Retrieved from http://jcmc.indiana.edu/vol3/issue1/paccagnella.html

Poole, N. (2013). Who do you know? A study of connectedness in online education and employment. *Education for Information*, *30*(3, 4).

Rettie, R. (2003). Connectedness, awareness and social presence. *PRESENCE 2003*, *6th Annual International Workshop on Presence.* Aalborg, Denmark. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.101.9820&rep=rep1&type=pdf

Richardson, J. C., & Swan, K. (2003). Examining social presence in online courses in relation to students' perceived learning and satisfaction. *Journal of Asynchronous Learning Networks, 7*(1), 68-88.

Rovai, A. P. (2007). Facilitating online discussions effectively. *Internet and Higher Education, 10*(1), 77-88. doi:10.1016/j.iheduc.2006.10.001

Rutter, J. and G. W. H. Smith (2005). Ethnographic presence in a nebulous setting. In C. Hine (Ed.), *Virtual methods: Issues in social reseach on the Internet* (pp. 81-92). New York, NY: Berg.

Rybas, N. and R. Gajjala (2007). Developing cyberethnographic research methods for understanding digitally mediated identities. *Forum: Qualitative Social Research 8*(3): 1-15.

Savolainen, R. (2015a). Approaching the affective factors of information seeking: the viewpoint of the information search process model. *Information Research*, *20*(1).

Savolainen, R. (2015b). Cognitive barriers to information seeking: A conceptual analysis. *Journal of Information Science*, *41*(5), 613-623.

Scollins-Mantha, B. (2008). Cultivating social presence in the online learning classroom: A literature review with recommendations for practice. *International Journal of Instructional Technology & Distance Learning*, *5*(3), 1-15.

Snowden, L. R. (1998). Racial differences in informal help seeking for mental health problems. *Journal of Community Psychology*, *26*(5), 429-438. DOI: 10.1002/(SICI)1520-6629(199809)26:5<429:AID-JCOP3>3.0.CO;2-M

Song, M., & Yuan, R. (2015). Beyond Social Presence: Increasing Cognitive Presence through Meaningful Interaction. *Global Learn,* 1, 731-736).

Spurgin, K. M. & Wildemuth, B. M. (2009). Content analysis. In B. M. Wildemuth (Ed.), *Applications of social research methods to questions in information and library science* (pp. 297-307). Westport, CT: Libraries Unlimited.

Stein, D., & Wanstreet, C. (2003). Role of social presence, choice of online or face-to-face group format, and satisfaction with perceived knowledge gained in a distance learning environment. *2003 Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education,* 193-198.

Suler, J. (2004). The online disinhibition effect. *Cyberpsychology & Behavior*, *7*(3), 321-326. doi:10.1089/1094931041291295

Teli, M., F. Pisanu, et al. (2007). The internet as a library-of-people: For a cyberethnography of onlne groups. *Forum: Qualitative Social Research 8*(3), 1-18.

Todd, R. J. (2005). Information intents. In K. E. Fisher, S. Erdelez & L. McKechnie (Eds.), *Theories of information behavior* (pp. 198-203) (ASIST Monograph Series). Medford, N.J: Information Today.

Trespalacios, J., & Rand, J. (2015). Using Asynchronous Activities to Promote Sense of Community and Learning in an Online Course. *International Journal of Online Pedagogy and Course Design (IJOPCD)*, *5*(4), 1-13.

Tu, C. H. (2000). On-line learning migration: From social learning theory to social presence theory in a CMC environment. *Journal of Network and Computer Applications, 23*(1), 27-37. doi:10.1006/jnca.1999.0099

Tu, C. H. & McIsaac, M. (2002). The relationship of social presence and interaction in online classes. *American Journal of Distance Education, 16*(3), 131-150. DOI: 10.1207/S15389286AJDE1603_2

Ward, K. (1999). Cyber-ethnography and the emergence of the virtually new community. *Journal of Information Technology, 14*(1), 95-105.

Wittel, A. (2011). Qualities of sharing and their transformations in the digital age. *International Review of Information Ethics*, *15*(09), 4-8.

Ybarra, M. L., & Suman, M. (2006). Help seeking behavior and the Internet: a national survey. *International Journal of Medical Informatics*, *75*(1), 29-41. doi:10.1016/j.ijmedinf.2005.07.029

Yu, L. (2012). Towards a reconceptualization of the 'information worlds of individuals'. *Journal of Librarianship and Information Science*, *44*(1), 3-18.

Yukawa, J. (2010). Communities of practice for blended learning: Toward an integrated model for LIS education. *Journal of Education for Library and Information Science*, 54-75.

Zimbardo, P.G. (1970). The human choice: Individuality, reason, and order versus deindividuation, impulse, and chaos. In W.J. Arnold & D. Levine (Eds.), *Nebraska Symposium on Motivation.* University of Nebraska Press. Lincoln, Nebraska.