MOOCs for Research: The Case of the
Indiana University Plagiarism Tutorials and Tests

Theodore Frick
Department of Instructional Systems Technology
School of Education
Indiana University Bloomington

Cesur Dagli
Department of Instructional Systems Technology
School of Education
Indiana University Bloomington

Revised Manuscript Accepted by Technology, Knowledge and Learning

June 14, 2016
MOOCs for Research: The Case of the
Indiana University Plagiarism Tutorials and Tests

Abstract

We illustrate a very recent research study that demonstrates the value of MOOCs (Massive Open Online Courses) as vehicles for research. We describe the development of the Indiana University Plagiarism Tutorials and Tests (IPTAT). Our new design has been guided by First Principles of Instruction: authentic problems, activation, demonstration, application, and integration. We further discuss our data collection mechanisms and early usage of this new mini-MOOC. In the first study, we investigated a built-in assessment feature for students to evaluate instructional quality and user experience. To do this, we adapted scales from the Teaching and Learning Quality (TALQ) instrument. As a follow-up study, we plan to further investigate patterns of usage of the IPTAT by students through creation of individual temporal maps. We plan to use Analysis of Patterns in Time, a method that provides learning analytics.

Keywords: educational research, MOOCs, mini-MOOCs, online instruction, online assessment, e-learning, student learning achievement, teaching and learning quality, temporal maps, Analysis of Patterns in Time, plagiarism, learning analytics.

(7,214 words in manuscript body, excluding references)
Introduction

In *Restructuring Education Through Technology*, Frick (1991) predicted that computers and the Internet would empower teachers and students in ways that could only be imagined at that time. Unknown to Frick, Tim Berners-Lee was at about the same time developing the HTML protocol for scientists to share their work via computers running server processes that would fetch and send documents across the Internet to other computers which had requested those files (Berners-Lee, 1990).

We now know this hyperlinked, document-sharing system as the World Wide Web. The Web has transformed ways in which we can communicate—and educate. This transformation is reminiscent of Gutenberg’s invention of the printing press and its widespread impact during and after the Renaissance in Western Europe over 500 years ago. Massive Open Online Courses (MOOCs) have emerged as educational vehicles that can empower teaching and learning worldwide.

In 2002, the first author and several graduate students developed a Web-based tutorial and test intended primarily for use by students in the Instructional Systems Technology program at Indiana University. Since this instruction and assessment was delivered via the Web, many others soon found it. Over the past 14 years, millions of visitors worldwide have benefitted from this online instruction and assessment. In actuality, this instruction was indeed a MOOC, or what Spector (2014) has recently referred to as a mini-MOOC—though not conceived as such when the plagiarism tutorial was originally developed.

About ten years ago, the first author began to consider the possibility of utilizing MOOCs as vehicles for data collection, so that members of his research groups could
carry out empirical studies. One of the issues many educational researchers often face is gaining access to a large population for gathering data. Historically, most research studies effectively have used relatively small *convenience* samples, often limiting generalizability of findings.

What if MOOCs were designed so data could be collected that could be used to answer research questions, particularly questions involving effectiveness of instructional methods and student learning achievement? Results of such studies could not only inform designers of MOOCs, but also could have broad implications for improving instruction to promote student learning achievement in more traditional face-to-face instructional contexts.

**Overview**

In this article, we first briefly discuss characteristics of MOOCs and major problems with MOOCs. We then describe *First Principles of Instruction* as a means of addressing deficiencies in MOOCs. Next we describe the Indiana University Plagiarism Tutorials and Tests (IPTAT), a MOOC originally developed in 2002. We used *First Principles* as guidelines for the major redesign and implementation of IPTAT in early 2016. We then describe two research studies, the first now completed, and the second planned for later in 2016. Both of these studies use IPTAT as a vehicle for research.

**Study 1.** This research study incorporated a way to collect user evaluation of IPTAT by using a modified version of the Teaching and Learning Quality Scales (TALQ). The TALQ was originally developed for use in evaluating formal courses in higher education—both face-to-face and through distance education (Frick, Chada, Watson, Wang & Green, 2009; Frick, Chada, Watson & Zlatkovska, 2010a; Frick,
Chadha, Watson & Yang, 2010b). The main focus was: Could an adapted version of the TALQ be used to reliably and validly evaluate the quality of MOOCs and mini-MOOCs?

**Study 2.** This study will be using a research method called Analysis of Patterns in Time (APT—Frick, 1990; Frick, Myers, Thompson, & York, 2008; Myers & Frick, 2015) in order to study instructional effectiveness. Students are in control of which parts of the IPTAT they undertake and complete. Behind the scenes, IPTAT software will create individual temporal maps that track each student’s choices as he or she navigates and completes learning activities in the IPTAT, including that student’s performance on a Certification Test. Patterns observed in temporal maps will be used to address the main research questions:

- When students experience fewer *First Principles of Instruction*, is the likelihood of their learning achievement lower?

- When students experience more *First Principles of Instruction*, is the likelihood of their learning achievement higher?

Finally, we briefly discuss future studies planned—all of which take advantage of temporal data collection strategies built into the design and implementation of IPTAT.

**What Are MOOCs and mini-MOOCs?**

The term, ‘MOOCs’, was created by Dave Cormier in 2008 to describe the first MOOC that was created by Siemens and Downes (Yuan & Powell, 2013)—a massive open online course. Wiley (2012) points out that although there are a number of online courses that are described as MOOCs, most do not follow at least one letter of the acronym. He agrees that most are online: however, not all MOOCs are “massive”, “open” or “courses.”
The first letter, ‘M’, of the acronym for the concept of scale refers to the number of participants (Koutropoulos & Hogue, 2012; Stewart, 2013). The discussion point here is how many participants should be in the course for it to be described as massive (Wiley, 2012). Anderson (2013) described massive as a general scale rather than a specific number of students. Also, massive may indicate a vast variety of resources and materials (Bond, 2013).

The second letter, ‘O’, in MOOC stands for ‘open’. Although there is an ongoing debate on what openness means in education (Peter & Deimann, 2013; Walker, 2005), ‘open’ in a MOOC refers to the free registration and participation from anywhere and to the access and use of course resources for free (Hylén 2006; Schaffert & Geser 2008).

The third letter, ‘O’, in MOOC denotes ‘online’. The word ‘online’ is the only term in MOOC about which many scholars have agreed on a common meaning, which refers to the delivery of courses through the Internet (Anderson, 2013). Online courses require students to have a device with Internet access in order to be able to participate.

The last letter of the acronym is ‘C’ meaning ‘course’. MOOCs in some ways differ from courses in a traditional educational setting (Wiley, 2012). Traditional courses are seen as being based on a particular subject with certain schedule times and run by an instructor who can deal with a manageable number of students (e.g., 25-100). Youell (2011, p. 4) defines a course as “a coherent academic engagement with a defined set of learning outcomes.” Courses in MOOCs do not necessarily require grading or assessment of student performance, whereas traditional courses do.

Spector (2014) has made a further distinction between MOOCs and mini-MOOCs. In his view, a mini-MOOC is a self-contained module of instruction that
largely stands on its own; contains built-in assessment and automated feedback for students; is shorter in length and scope than is a course; is specifically focused on a narrow instructional objective; does not have a specified human instructor available to work with students—instead authentic activities can be provided; nor is it as comprehensive as what we normally conceive of as a course for credit in higher education.

**Major problems with MOOCs**

Although the number of people who pursue learning in MOOCs has been growing (Bull, 2012; Lombardi, 2013), little is known about instructional methods and strategies that effectively promote successful e-learning in MOOCs (Bujack, Paul & Sandulli, 2012). Researchers have identified that deficiencies exist with regard to designing, developing, and deploying MOOCs (Downes, 2013a; Spector, 2014; Wiley, 2007; Williams & Su, 2015). Because they are self-contained content-delivery based learning packages, the quality of course content and online delivery methods such as videos and audios take particular importance in MOOCs, especially in xMOOCs (Venkataraman & Kanwar, 2015; Walker & Loch, 2014).

MOOCs have also been criticized for lack of sufficient feedback and weak forms of assessment, which are considered to be main reasons for low engagement and completion rates (Bates, 2012; Bonk, Reeves, Reynolds, & Lee, in review; Daniel, 2012; Rodriguez, 2012; Yang, Wen & Rose, 2014). The needs of MOOCs students, particularly xMOOCs students, who have limited or no access to their instructors, should also be taken into consideration in the design process. These current deficiencies in the design and implementation of MOOCs may be addressed by closer adherence to instructional
design (ID) theories (Kopp, & Lackner, 2014; Bonk, Lee, Reeves, & Reynolds, 2015; Margaryan, Bianco & Littlejohn, 2015).

Reigeluth (1999) emphasized that ID theories should propose prescriptive guidelines, including methods and principles, for designing effective instruction for various contexts. Reigeluth and Frick (1999) further stated:

At the very least, [ID theories] can all benefit from more detailed guidance for applying their methods to diverse situations. And more theories are sorely needed to provide guidance for additional kinds of learning and human development and for different kinds of situations, including the use of new information technologies as tools (p. 633).

Merrill (2002) proposed a set of interrelated prescriptive instructional principles synthesized from a number of ID theories and models, which he called First Principles of Instruction. Merrill claimed that the effectiveness of instruction depends upon the extent to which these principles are successfully implemented, regardless of learning environments, programs, and practices. This claim has been empirically investigated in a few studies (see Merrill, 2013, Chapter 22), but not in the context of MOOCs or mini-MOOCs. Therefore, the instructional principles proposed by Merrill may address deficiencies in the design of MOOCs.

**First Principles of Instruction**

Merrill (2002, 2009) proposed a set of interrelated prescriptive instructional principles synthesized from a number of ID theories and models: task/problem-centered learning activities, activation, demonstration, application, and integration, which he called “first principles of instruction”. According to Merrill (2013), a principle is “a relationship that is always true under appropriate conditions, regardless of the methods or models used to implement the principle” (p. 19). To call a relationship a principle, it must be verified by
empirical research, which demonstrates that the principle is effective, efficient, or engaging, and the principle is applicable in any learning setting (Clark, 2008). Merrill (2002) stated the properties of *First Principles of Instruction* as follows:

First, learning from a given program will be promoted in direct proportion to its implementation of first principles. Second, first principles of instruction can be implemented in any delivery system or using any instructional architecture. Third, first principles of instruction are design oriented or prescriptive rather than learning oriented or descriptive (p. 44).

Therefore, Merrill (2002) claimed that the effectiveness of instruction depends upon the extent to which the principles are successfully implemented—regardless of learning environments, programs, and practices. *First Principles of Instruction* aim to aid in the design and creation of instructional products and learning environments, but they do not describe how learning occurs in learning settings. *First Principles of Instruction* specify a four-phase cycle of instruction that is centered around a series of increasingly complex authentic problems or tasks. Each of these learning cycles should include activation, demonstration, application, and integration (Merrill, 2002).

**Problem-centered: Let me do the whole task!**

The initial first principle is that instruction should be problem- or task-centered. Merrill proposed that solving real-world problems helps promote learning. After instruction is complete, students should be able to perform the whole task or solve the whole problem. The main focus is to begin with an authentic, real-world problem, and then to sequence problems from simple to complex.

According to this ID theory, a problem is a wide range of activities that can be related to whole tasks that students will face in the real word. Utilizing knowledge and skills out of context often prevents their retrieval when needed. In other words, de-
contextualization of instructional components can impede performance of desired skills in contexts of real-world situations. “Learning is promoted when learners engage in a problem-centered instructional strategy in which component skills are taught in the context of a simple-to-complex progression of whole real-world problems” (Merrill, 2013, p. 27).

**Activation: Where do I start?**

“Learning is promoted when learners activate a mental model of their prior knowledge and skills as a foundation for new skills” (Merrill, 2013, p. 28). The focus in the activation phase is to stimulate students’ previous knowledge and experience with their existing mental schema—as a foundation for acquiring new skills and knowledge. Activating previous student experience helps minimize frustration, prepares them for new learning, and can increase instructional efficiency. If students lack prior knowledge or skills, this phase highlights provision of relevant experience prior to new learning. Organizing and structuring relevant themes of content being taught can provide mental models and help students connect new mental structures with their existing mental structures.

**Demonstration: Don’t just tell me. Show me!**

The demonstration principle emphasizes that instruction should provide students with examples of what is to be learned. This principle focuses on the importance of specific demonstration cases or situations to gain knowledge rather than providing general information through tell-and-ask instruction only. The consistency of demonstration with desired student learning outcomes helps make instruction more effective. Demonstration should focus on structural features of problems with multiple
and different representations. Appropriate combinations of text, graphics, and audio in media can reduce the cognitive load of students and promote learning. “Learning is promoted when learners observe a demonstration of skills to be learned” (Merrill, 2013, p. 23).

**Application: Let me do it!**

After activating prior knowledge and demonstrating what to be taught, the next phase is for student performance of desired skills or use of new information. In the application phase, Merrill (2002) emphasized applying new knowledge that is consistent with the content being taught in order to promote learning. Providing consistent practice for a learning goal is most important in this phase. Merrill emphasized that instruction should gradually decrease coaching and support as learning progresses. Feedback on task practice allows students to recognize and correct their mistakes and helps enable improved performance. Applying knowledge to a variety of problems rather than a single problem increases the likelihood of successful learning. “Learning is promoted when learners apply their newly acquired knowledge and skills” (Merrill, 2013, p. 25).

**Integration: Watch me!**

This principle advocates having students utilize their new knowledge in real world situations. Exhibition to others of what students have learned and recognition of their learning progress can be motivating factors. Other ways to facilitate integration include reflecting on, defending, and sharing newly acquired skills and knowledge. Personalizing and taking ownership of what a student has learned helps to reinforce new mental structures, according to Merrill. “Learning is promoted when learners reflect on, discuss, and defend their newly acquired knowledge and skills” (Merrill, 2013, p. 29).
The Original IPTAT: 2002 – 2015

In our view, the IPTAT developed originally in 2002 best fits Spector’s (2014) conception of a mini-MOOC. It should be noted that the IPTAT was developed long before MOOCs became identified as such.

The tutorial and test on How to Recognize Plagiarism was originally developed for use by students in the Instructional Systems Technology (IST) department at Indiana University, starting in September, 2002.

Figure 1. Home page of the original tutorial, circa 2003.

As can be seen in the sidebar in Figure 1, the original IPTAT included plagiarism cases, examples, practice, and a test. In Figure 2, a part of the 10-item test is illustrated.
As others have since discovered these online resources on the web and by word of mouth, usage has been increasing almost exponentially each year. See Figure 3.
Throughout the 14 years of design, development, modification, and usability testing, the plagiarism tutorial design team has consisted of a variety of members. See https://www.indiana.edu/~academy/firstPrinciples/credits.html. The design team has learned through correspondence initiated by instructors that many now require their students to pass this test. We have no control over who uses our tests and for what purposes. Our goal is to help people understand what plagiarism is, so that they do not commit plagiarism in their writing and presentations. Anyone is welcome to use our learning resources and tests for free.

Aside from minor corrections and modifications, the original tutorial and 10-item test remained largely the same between 2002 and 2012. In 2013, several instructors e-mailed us because they were very concerned about cheating on the Certification Test. A YouTube video had been posted with the answer key, viewed by thousands. After several rounds of reordering questions, followed by new corresponding answer keys that were posted soon afterwards, we decided to take a different approach.

Large test item pools were developed, one primarily for undergraduate students and the second for graduate students. The latter item pool consisted of more difficult questions, with more subtle and well-disguised examples of plagiarism. Each 10-item certification test consisted of randomly selected questions, resulting in billions of unique tests for undergraduates and trillions for graduate students.
These item pools were respectively implemented in late 2013 and early 2014. For further details on changes the design team made in order to minimize cheating, see: https://www.indiana.edu/~academy/firstPrinciples/recentChanges.html.

**Redesign of the IU Plagiarism Tutorials and Tests (IPTAT)**

We redesigned the IPTAT for several reasons: First and foremost, we wanted the new design to incorporate *First Principles of Instruction* (Merrill, 2002, 2013). We did this specifically so that we could carry out research studies to evaluate how these five principles of instruction affect student learning.

Second, and also highly important, we wanted to improve the effectiveness of our instruction in the mini-MOOC. The original tutorial, developed in 2002, did not specifically follow *First Principles of Instruction* to guide its development, since we did not know about Merrill’s research (published later in 2002). Also, we noted in recent years that many students had difficulty in passing Certification Tests, typically taking between 7 and 10 attempts to pass in 2013-15. This problem was inadvertently created, in part, because Certification Tests had been significantly changed in 2013-14 order to make it harder to cheat. Some students told us directly that the instruction was not adequate preparation for passing a test, and they were right.

Third, we wanted to incorporate multimedia where it made sense. This was not practical in 2002 due to limited bandwidth, when a large portion of users had dial-up connectivity. Furthermore, the IPTAT should be scalable for a range of display sizes—from smartphones and tablets to laptop and desktop computers.
We chose not to provide instructor-to-student interaction nor student-to-student interaction in the new IPTAT. Because of the very large scale of usage (literally millions of students worldwide) and very limited resources, those features that are often part of MOOCs were excluded from the new IPTAT design. Our goal was to design the mini-MOOC so that it literally runs by itself, requiring minimal management by humans—other than occasional monitoring of usage logs and reading feedback and comments provided by students and instructors. The IPTAT is truly open, and free to use by anyone. In the past 14 years, design, production, software engineering, maintenance, and management of the IPTAT has been done by volunteers who were not paid for this work.

The redesign process took place over a period of about 9 months, with the bulk of the development and production completed in late 2015. Space does not permit description of this design process and critical design decisions we made along the way that affected the final design. What we describe below is the production version of the IPTAT now available at https://www.indiana.edu/~academy/firstPrinciples/. A summary of differences between the original and the new design are provided at: https://www.indiana.edu/~academy/firstPrinciples/choice.html.

**Authentic problems principle**

This required us to design a series of authentic problems in recognizing plagiarism, arranged from simple to complex. We did so, as reflected on menu at: https://www.indiana.edu/~academy/firstPrinciples/tutorials/index.html. As can be seen in Figure 4, problems are arranged at 5 levels of difficulty in recognizing plagiarism: basic, novice, intermediate, advanced, and expert. At each level of difficulty, we provide activation, demonstration, application, integration, and a practice test.
Figure 4. Five levels of difficulty in recognizing plagiarism.

Activation principle

We decided to design and develop 10 video cases as a means of student activation—providing real-world cases for students to experience vicariously. Storytelling is an instructional method which we implemented in our new design to embody the activation principle (e.g., see Andrews, Hull & Donahue, 2009). View an example of a video case at:

https://www.indiana.edu/~academy/firstPrinciples/tutorials/task1/activation.html

Similar video cases that tell stories are provided at each of the 5 levels of task difficulty.
Figure 5. Initial video case in the IPTAT on the starting page, for the activation principle.

Demonstration principle

Here we chose to design 12 screencasts that dynamically portray the writing of a whole, short paper (with voice over). Examples dynamically show the author committing plagiarism and how he fixes it in order to avoid plagiarism. See, for example, demonstrations at task level 2:

Application principle

Here we developed questions similar to those on the Certification Tests, but with two differences. Question difficulty matches the level of task difficulty at each of the 5 levels. Immediate feedback on the correctness of each answer is provided. In addition, if the answer is incorrect, a detailed explanation of why it is incorrect is provided. Also if needed, explanation is provided on how to correctly fix the student version to avoid plagiarism. For example, see task level 3 practice items:

Integration principle

This was perhaps the most challenging principle of instruction to implement in a mini-MOOC. We decided to do this by giving students an opportunity to reflect on what they just learned and how it might be used in their own lives. See for example the integration activity at the task 4 level of difficulty:

https://www.indiana.edu/~academy/firstPrinciples/tutorials/task4/integration.html. Here we do not provide feedback on what students write in the text input box, but we do store
their comments for later qualitative content analysis to be done as part of research studies.

![Instruction: Reflect on what you've learned at the Advanced Level](image)

Figure 8. Example of a reflection activity, for the integration principle.

**Practice tests for self-assessment of learning progress**

Each practice test mirrors the format of a Certification Test but questions are provided which match the level of task difficulty. Practice test items at the expert level provide the most difficult questions, for example: [https://www.indiana.edu/~academy/firstPrinciples/tutorials/task5/masteryTest.php](https://www.indiana.edu/~academy/firstPrinciples/tutorials/task5/masteryTest.php). After answering all questions, explanatory feedback is provided on each answer. Practice tests differ from Certification Tests in that the latter provide only feedback on the kinds of mistakes being made, not on specific questions missed or correct answers to those questions.

**Certification Tests for assessing mastery of IPTAT learning objectives**

We have designed large inventories of test items, respectively, for (a) undergraduate and advanced high school students, and (b) for master’s and doctoral students. As might be expected, test questions are generally much more difficult for the latter group. Since questions are selected at random from large inventories, each 10-item test typically contains questions from the 5 task levels. What makes questions harder at
the graduate level is that the plagiarism is more skillfully disguised than at the
undergraduate level. In order to take a Certification Test, each user must register by
providing their unique e-mail address and a password they create. After verification by
clicking on a link in the e-mail message sent to them, their registration is activated and
they can take a test. See for example:


**Demographic survey**

When users first register, they complete a brief demographic survey. There are 4
optional questions on their: age, highest level of education achieved, reason for doing the
IPTAT, and self-appraisal of current level of understanding of plagiarism. See for
example:

https://www.indiana.edu/~academy/firstPrinciples/login.phtml?action=register&testLevel=UG.

**MOO-TALQ survey**

We have embedded this survey in the IPTAT as an optional evaluation activity for
users prior to taking a Certification Test. The MOO-TALQ is described below in more
detail in the methods section for Research Study 1.

Readers are invited to try the IPTAT themselves and take tests in order to become
acquainted with the details of this mini-MOOC and the MOO-TALQ questions.

**How data are collected in the IPTAT and who is using it?**

The IPTAT runs on a Web server at Indiana University. The current site has over
500 HTML files used for this mini-MOOC. For parts of the instruction, surveys, practice
questions, and Certification Tests, we have done extensive PHP programming, which in
MOOCs for Research – 23

turn dynamically generates HTML code that is delivered to a user’s Web browser for
interpretation and display.

PHP software includes code that stores data in two formats in two different
places, neither of which can be accessed through Web browsers, in order to maintain data
security. Details about our schemes cannot be described here for security reasons, so we
can only be general. In one location, sequential ASCII files are created and records are
appended as needed. These kinds of files are easily readable by humans. For example, a
line in our file for recording whenever someone passes a Certification Test looks like this
(the actual e-mail address and user name have been modified here to protect his/her
identity):

xyz@gmail.com | Name | 202.120.38.125 | 201690130281137511586 | 2016-02-28 10:15:12 | 7 | GR

In this example, we can further see the user’s IP address, test ID, date/time, minutes taken
to pass, and that it was a graduate-level test.

Note that each record (a line in the file) includes the unique user e-mail address,
which is a primary key that can be used to cross-link other ASCII files which record test
attempts, MOO-TALQ results and, eventually, within temporal maps for tracking what
each user does (to be created in Study 2, later in 2016).

The second data storage mechanism is a separate MySQL database, where the
primary key is also each user’s e-mail address. The MySQL database table contains one
row for each user, and has columns for user registration information, activation,
demographic survey results, and information about the last Certification Test passed. The
MySQL server runs on a different host computer than does the Web server. This
provides some redundancy with the sequential ASCII file system, but in a separate place
to decrease vulnerability. Both data storage systems are backed up daily to other locations to help prevent accidental data loss.

For the new version of the IPTAT, we are using Google Analytics to track and report usage. For example, in the first 5 months in 2016 after the new IPTAT was implemented, we have data as follows:

- 283,000+ user sessions, each about 24 minutes long, with 46% return visitors, from 179 countries worldwide, about 94% of whom are from the U.S., China, and Canada.

- About 5.6 million page views of which 3.4 million are unique.

- About 180,000 video plays that have been viewed to completion (our video cases and demonstration screencasts).

- At peak usage times there are typically 175 – 200 concurrent user sessions occurring in real time. Peak usage times are typically Monday through Friday, and higher during the beginning of the semester (e.g., the month of January).

From our MySQL database, we know that between January 2 and June 8, 2016, we have had over 74,000 users successfully register by activating via their e-mail; over 49,000 have passed a Certification Test for undergraduates or high school students; and over 10,000 graduate students have passed.

Clearly, we have massive usage of our mini-MOOC and large numbers of cases in order to pursue research studies described below.
Study 1: MOO-TALQ as an Evaluation Instrument for Online Instruction

Research questions

The main research questions in Study 1 were:

1. What are the relationships among student ratings on MOO-TALQ scales and student mastery of how to recognize plagiarism?

2. If students agree that they experienced First Principles of Instruction and Academic Learning Time in the IPTAT, is the likelihood of their mastery greater than that of students who do not agree?

Academic Learning Time (ALT) refers to the frequency and time students are successfully engaged in learning tasks that are aligned with assessments of learning achievement (e.g., Rangel & Berliner, 2007; Frick, et al., 2010b).

Methods

Measurement instruments. In order to evaluate learning experiences in MOOCs and mini-MOOCs, a new instrument was created, called MOO-TALQ. MOO-TALQ scales were adapted from those from the teaching and learning quality (TALQ) survey (Frick et al., 2009, 2010a, 2010b). They constructed a course evaluation instrument that provides for student ratings of First Principles of Instruction and Academic Learning Time (ALT). The TALQ scales were found dependable in both face-to-face and online learning environments (Chadha, 2009). The TALQ scales can be utilized to “indicate areas of improvement needed in existing courses” (Frick, Koh & Chadha, 2011, p. 44).

MOO-TALQ scales were created through modification of the TALQ instrument. The original TALQ consists of 40 items with nine a priori student rating scales: Academic Learning Time, learning progress, student satisfaction, global quality, authentic
problems, activation, demonstration, application, and integration (Frick et al., 2010a). In comparison, the MOO-TALQ was designed to evaluate learning experiences in massively open online contexts (what Spector, 2014, referred to as mini-MOOCs). The new MOO-TALQ consisted of 25 items. There were two main reasons for decreasing the number of items in the original TALQ survey from 40 to 25. First, in his dissertation study, Chadha (2009) examined the dependability of TALQ scale scores and suggested changes in the TALQ survey. In addition, the fact that students taking this tutorial intend to pass a Certification Test as soon as possible was also an important consideration for reducing the number of items in the survey in order to encourage participants to complete the survey (Frick et al., 2014).

The eight scales in the new MOO-TALQ are designed for rating overall quality and satisfaction; task engagement; task success; authentic problems; activation; demonstration; application; and integration. Each scale consists of two to five items and employs a five-point Likert scale. The Likert responses indicate a degree of agreement or disagreement with the statement (strongly disagree = 1, disagree = 2, undecided = 3, agree = 4, strongly agree = 5), and the sixth option provides for a rating of ‘not applicable’. The Likert items are randomly mixed, not organized according to the MOO-TALQ scales. Reliabilities of subscales in the MOO-TALQ instrument ranged from 0.74 to 0.94 as indicated by Cronbach’s α coefficient for internal consistency.

*Student learning achievement.* Student mastery was assessed by criterion-referenced Certification Tests. There were two separate Certification Tests: one for high school and undergraduate (H&UG) students, and one for master’s and doctoral graduate (GR) students. Each individual test consisted of a set of 10 randomly selected questions
from very large item pools, respectively, with billions and trillions of unique tests possible. Students who passed a test were awarded a Plagiarism Certificate. They were later classified as high and low masters at the 95% confidence level, and those remaining were classified as medium masters. This was done by use of Bayesian Reasoning and the Sequential Probability Ratio Test (Wald, 1947). See Dagli (2016) for further details.

*Participant selection criteria and data collection.* Data were collected between January 12 and 31, 2016, inclusively. During this 20-day interval, 36,801 users registered to take a test, including study participants and non-participants.

Eligibility criteria for voluntary participation in Study 1 included: (1) agreement to participate in the study, (2) completion of the MOO-TALQ survey, (3) completion of one Certification Test, and (4) use of only the first attempt for both the survey and test for data analysis. After applying participant selection criteria to identify qualified cases, the resulting number of participants in this study was 2,016. Overall, 85% of eligible study participants took H&UG Certification Tests; and 15% took GR Certification Tests.

*Participant demographics.* Approximately 54% of the participants reported a high school diploma as their highest degree, 19% reported an associate’s degree, 15% reported a bachelor’s degree, 4% reported a master’s degree, and 0.6% reported a doctoral degree. At the time of registration, nearly 50% of these users rated their current level of understanding of plagiarism as confident, 48% as some confidence, 1% as little confidence, and 1% did not answer this question. The majority (59%) were between 18 and 22 years old, 35% older than 22, and 6% younger than 18. Approximately 96% indicated they took certification tests to fulfill a course requirement, and 2% reported that they wanted to learn how to recognize plagiarism. Of the 1,716 participants who took
Certification Tests for undergraduate and advanced high school students, 85% passed a Certification Test, whereas of the 300 participants who took Certification Tests for graduate students, 86% passed a test.

**Results**

*Research question 1.* Spearman correlation coefficients were calculated to indicate associations between student mastery level and ratings of their perceived experience of *First Principles of Instruction*, as well as their perceptions of *Academic Learning Time* (ALT), and rating of overall satisfaction with IPTAT. Each MOO-TALQ scale consisted of Likert ratings of two or more individual items that comprise that particular scale. Correlational analyses of the MOO-TALQ ratings indicated that there were strong positive correlations in both groups between ratings of *First Principles*, ALT, and overall quality of the MOOCs and student satisfaction with what they learned. Spearman correlations among MOO-TALQ scales were very high, ranging from 0.558 to 0.855. Student ratings on each scale averaged about 4, meaning that they ‘agreed’ with scale items—that is, that the IPTAT was of high quality and that they were successfully engaged.

There were weak though highly statistically significant bi-variate relationships between student mastery and *First Principles of Instruction* in the both H&UG and GR groups. The tri-variate relationship between *First Principles of Instruction*, ALT, and student mastery was better answered by APT and HLM analyses to answer question 2.

*Research question 2.* Contingency tables were further constructed; and Analysis of Patterns in Time (APT) was used to determine conditional probabilities of mastery levels for those who agreed versus those who disagreed that they experienced *First
Principles of Instruction. In addition, a hierarchical log-linear model (HLM) analysis was conducted as an alternative way to interpret more than two-way interactions, and to determine associations among the MOO-TALQ scales and student mastery.

APT analyses revealed clearer relationships between student mastery and First Principles, when compared with the Spearman correlational analyses. In the graduate student (GR) group, participants who agreed that they experienced First Principles of Instruction and ALT were about five times more likely to be high masters, when compared to those who did not agree that they experienced First Principles and ALT. Similarly, participants in the high school and undergraduate (H&UG) student group who agreed that they experienced First Principles of Instruction and ALT were about three times more likely to achieve high mastery than did those who did not agree. Moreover, based on the relative contribution of each principle, the H&UG test takers relied more on the demonstration and application principles to achieve mastery, while the GR students tended to experience all First Principles of Instruction; in addition, GR students rated their experience of the authentic problems principle and the integration principle more highly than other First Principles for achieving mastery. Findings from the APT analyses were further validated by findings from HLM analyses of both groups, which indicated significant interactions between mastery level and First Principles, and between First Principles and ALT.

Implications

These findings are consistent with Merrill’s claims and his prediction that First Principles of Instruction promote learning. Results from Study 1 indicate that employing First Principles of Instruction in the design of MOOCs is likely to yield high quality
instruction and satisfaction with MOOCs, as well as to promote what students learn within MOOCs. Furthermore, this study indicated that the MOO-TALQ scales are reliable and can provide an efficient way to evaluate learning experiences in the context of MOOCs and mini-MOOCs. For further details on Study 1, see Dagli (2016).

**Study 2: Analysis of Patterns in Time to Address the Effectiveness of First Principles of Instruction**

We are further investigating the effectiveness of *First Principles of Instruction* embodied in the new IPTAT. As explained above, the newly designed IPTAT has been guided by *First Principles of Instruction* (Merrill, 2013).

**Research Questions in Study 2:**

To the extent students are exposed to *First Principles of Instruction* in IPTAT:

- When fewer *First Principles of Instruction* are experienced, is the likelihood of student learning achievement lower?
- When more *First Principles of Instruction* are experienced, is the likelihood of student learning achievement higher?

**Methods**

In Study 1, student ratings on the MOO-TALQ were used to investigate *First Principles*. In Study 2, we will not rely on student self-reports. Instead, we will track the sequence of event occurrences about what each student actually does when interacting with IPTAT by creating a temporal map for him or her.

We plan to use a methodology called Analysis of Patterns in Time. APT is part of MAPSAT research methodology (Map & Analyze Patterns & Structures Across...
MOOCs for Research – 31

Time). MAPSAT differs from traditional quantitative educational research methods, where variables are measured separately and then relations among variables are analyzed statistically. In MAPSAT, relations themselves are empirically observed and coded. MAPSAT was invented decades ago, and is well suited as a methodology for modern learning analytics, as well as for many other kinds of research (Frick, 1990; Myers & Frick, 2015). For example, Google Analytics utilizes a variation of APT methods in its “behavior flow” reports; however, temporal maps cannot be queried as in full APT.

In APT, measures of relations are determined by relative frequency and/or duration of occurrences of observed temporal patterns. In other words, researchers code sequences of occurrences of events using defined categories from multiple classifications in an observation system. This results in a temporal map for each unique observed entity (i.e., each student who tries to learn via IPTAT). We are further developing software that will track each student’s use and store it as a temporal map. See Myers and Frick (2015) for examples of how we have used temporal maps with APT in studying play-learner usage of the online Diffusion Simulation Game.

Each temporal map in APT can be represented by a spreadsheet. The rows in the spreadsheet represent successive moments in time; the columns represent the classifications in the observation system; and category names are entered by an observer into spreadsheet cells (in Study 2, this will be done with software). The entries into the cells represent the temporal order of specific empirical events which are observed to occur within each classification column. The rows in the spreadsheet are labeled by the date and time of each event occurrence. After observations are completed, a researcher subsequently can count specific qualitative patterns within each unique temporal map, as...
well as sum the durations of a particular temporal pattern. See a general example of APT at: http://educology.indiana.edu/affectRelationTemporal.html.

For example, in the study of First Principles of Instruction, we plan to conduct APT queries that count patterns representing the sequence of specific instructional principles that, in turn, are followed by student mastery of learning objectives. There will be a temporal map for each student who goes through the online IPTAT. Observations of event occurrences will be done by computer software embedded in the online instruction which will be using specific codes that the researchers have previously associated with each activity (e.g., this activity is an instance of the application principle, or that activity is an instance of the activation principle). When a student takes a test, computer software will classify the student as a master or nonmaster of the learning objective. This will result in literally thousands of temporal maps. Probabilities of event sequences leading to student mastery can be estimated by APT queries of temporal maps. Each query will scan the temporal maps for occurrences of temporal patterns and count them.

In summary, the most salient difference is that MAPSAT measures relations, whereas quantitative statistical methods relate measures. This is not a play on words, rather a profound difference in approach to measurement and analysis in empirical research studies.

Results of APT queries about temporal patterns can be subsequently analyzed with traditional statistical methods. Results of an APT query about a specific temporal pattern can be treated as single variable which is a relationship itself, the value of which is the measure of that pattern for each temporal map. That is, the relationship has a value itself, in contrast with a relationship that is estimated statistically by a linear model such
as multiple regression or ANOVA. Such a *relationship measure* can be counts of the corresponding pattern occurrence in a temporal map, duration of that pattern occurrence, or likelihoods of that pattern occurrence. Such a *relationship measure* of pattern variable can, in aggregate, be treated statistically through means (averages), standard deviations, conditional probability estimates, confidence intervals, etc. For further details about temporal maps and APT queries, see Myers and Frick (2015).

**Further Planned Studies**

In Study 3, we will have an opportunity to investigate isomorphism between student ratings in MOO-TALQ and corresponding APT relationship measures from temporal maps that indicate what students actually *do* during the IPTAT. ‘Isomorphism’ is a structural measure for comparing systems (see Thompson, 2008). If findings indicate a high degree of isomorphism between these two measurement systems, then this would have significant implications for designers of MOOCs and mini-MOOCs. The practical implication is that, if MOO-TALQ ratings can strongly predict APT results, then the former would be a highly cost-effective way to evaluate instructional quality of MOOCs, and would not require inclusion of sophisticated software for storing, retrieving and analyzing temporal maps with APT in other MOOCs.

A fourth planned study may compare MOOCs that do and do not incorporate *First Principles of Instruction*, possibly via experimental treatments with various levels of *First Principles* present.
Summary

In conclusion, we have illustrated in this report how MOOCs can be used as vehicles for conducting educational research. When MOOCs are used by a large population, as we have had with the Indiana University Plagiarism Tutorials and Tests, results can have high generalizability. In fact, with such large numbers of cases, statistical significance becomes far less important than practical significance of findings on ways to improve instruction that, in turn, promote student learning achievement.

References


Bond, P. (2013). Massive open online courses (MOOCs) for professional development and growth. In C. Smallwood, K. Harrod, K. & V. Gubnitskaia (Eds.), Continuing
education for librarians: Essays on career improvement through classes, workshops, conferences and more (pp. 28-35). Jefferson, NC: Mcfarland & Co. Inc.


