Rebecca M. Quintana and Yuanru Tan University of Michigan

Abstract

We explore new tools and methods for learning designers and researchers to characterize pedagogical approaches that are applied to the design of MOOCs. This paper makes three main contributions to literature on MOOC design and evaluation: (1) an Expanded Assessing MOOC Pedagogies instrument for use by learning designers and researchers within their own contexts, (2) a demonstration of how nearest neighbor cluster analysis can be used to identify pedagogically similar MOOCs, and (3) a preliminary analysis of the clusters to account for features and factors that contribute to pedagogical similarity of MOOCs within clusters. This work advances research in the development of MOOC typologies, to allow learning designers and researchers to ask nuanced questions about pedagogical aspects of MOOC design.

Keywords: massive open online courses (MOOCs), assessment instrument, pedagogy, clustering methods

Quintana, R.M., & Tan, Y. (2019). Characterizing MOOC pedagogies: Exploring tools and methods for learning designers and researchers. *Online Learning*, 23(4), 62-84. doi:10.24059/olj.v23i4.2084

Characterizing MOOC Pedagogies: Exploring Tools and Methods for Learning Designers and Researchers

For many researchers and designers who analyze massive open online courses (MOOCs), it is essential to understand the pedagogical perspectives that are instantiated in the design of a course. Despite outward appearances, MOOCs are not a "single monolithic entity" (Major & Blackmon, 2016, p. 12), and great variation exists within the form. The constraints on design imposed by course delivery platforms are well known (Head, 2017), yet these restrictions do not necessarily result in a uniform approach to design. Many instructional teams have resisted a templated approach to design (Seaton, 2016) and have leveraged platform features to develop learning experiences that are mapped to learning goals (Najafi, Rolheiser, Harrison, & Håklev, 2015). Therefore, since MOOC designs and the pedagogical approaches that underlie them are divergent (Admiraal, Huisman, & van de Ven, 2014; Quintana, Tan, & Korf, 2018), it is important for learning designers and researchers to be able to meaningfully characterize these pedagogies. As Swan, Day, Bogle, and van Prooyen (2014) advocated, "finding mechanisms to distinguish among MOOCs or evaluate their underlying components or characteristics should be the first step in the research" (p. 75). By exploring the range of pedagogies that exist within MOOCs, learning designers and researchers can better understand design outcomes, including (1) assessing

alignment with design goals and outcomes and (2) understanding how one course compares to another (e.g., when multiple courses exist together within a series).

Yet, developing methods to systematically articulate learning design similarities and differences within MOOCs is challenging. Some MOOC researchers have endeavored to provide a holistic characterization of the instructional design of a course by studying its composition—the type, frequency, and arrangement of course elements (Quintana, Tan, & Korf, 2018; Seaton, 2016). Researchers identified various structures through a process of abstraction and relied on the separation of course content and its internal structure (Davis, Seaton, Hauff, & Houben, 2018). This approach allows designers and researchers to make high-level observations but does not offer insight into underlying approaches to design. Other researchers have sought to distinguish between types of MOOCs, making determinations about how they are similar and different from each other pedagogically (Major & Blackmon, 2016). The development of the Assessing MOOC Pedagogies (AMP) instrument represents a major advance in the study of MOOC pedagogies (Swan, Day, Bogle, & van Prooyen, 2014; Swan, Day, Bogle, & van Prooyen, 2015; Swan, Day, & Bogle, 2016). The AMP instrument was designed to characterize the pedagogical design that underlies a MOOC across 10 relevant dimensions (Swan et al., 2015). This approach is concerned with aspects that directly affect learning, such as approach to content presentation and the role of the learner within the course (Reeves, 1996). The AMP tool demonstrates potential to allow its users to make comparisons across multiple MOOCs. For instance, courses that exhibit similar dimensions could be grouped together, which could help researchers articulate a range of pedagogical typologies. Such progress could aid in MOOC design and evaluation and allow researchers to "ask meaningful questions" of these courses (Major & Blackmon, 2016, p. 20).

However, there are remaining challenges—the approaches that are outlined in the AMP literature (e.g., Swan et al., 2015) do not specify precisely *how* learning designers and researchers can make use of the instrument. There are still important details lacking, such as clear explanations of what differentiates one score from another within each dimension of the instrument. Using the present version of the instrument, reviewers would be hard-pressed to achieve consistency in scoring, leading to differences in understanding of a MOOC's design and low interrater reliability (IRR; i.e., the degree of agreement among multiple raters). Additionally, although the AMP literature indicates that pedagogically similar courses can be grouped once they have been scored using the instrument, there are also no clear guidelines on how to do this in new contexts (i.e., outside of the original studies).

Objectives

What is needed are nuanced expansions to the AMP instrument to provide more specific methods for understanding similarities and differences in design in an effort to support MOOC design and evaluation activities. Our overarching goal was to develop tools and methods for characterizing the pedagogical design of MOOCs, primarily to support learning designers and researchers who want to understand and articulate the pedagogical dimensions that are represented in MOOCs.

To achieve our overarching goal, we identified three subgoals:

• **Expand the AMP instrument.** Our goal was to explore how Swan et al.'s (2015) AMP instrument can be more easily adopted by researchers and practitioners through elaboration and potential modification.

- Use the Expanded AMP instrument to characterize the design of a MOOC. Our goal was to employ the updated AMP instrument to assess the pedagogical design of a set of MOOCs, toward identifying clusters of pedagogically similar courses.
- Use the Expanded AMP instrument for larger MOOC comparison. Our goal was to use the Expanded AMP instrument to develop and analyze MOOC clusters to understand how underlying features and factors make one group of courses different from another.

Given these objectives, we articulated the following research questions:

- 1. What elaboration and/or modifications are needed for learning designers and researchers to use Swan et al.'s (2015) AMP instrument to characterize the pedagogical design of MOOCs?
- 2. What features do pedagogically similar MOOCs share? What factors may contribute to this congruence?

Perspectives

The AMP instrument was initially developed to examine MOOCs that were being considered for college credit (Swan et al., 2015; Swan, Day, & Bogle, 2016). It was based on a pedagogical assessment tool for computer-based education developed by Reeves (1996). Swan et al. (2015) retained six of 14 dimensions: epistemology, role of the teacher, experiential validity (renamed "focus of activities"), cooperative learning, accommodation of individual differences, and user role. Swan et al. (2015) added four new dimensions to the AMP instrument that are relevant to the MOOC context: structure, approach to content, feedback, and activities and assessments. Swan, Day, Bogle, and van Prooyen (2014) explained that the AMP instrument is not intended to provide a rating of the quality of a MOOC but rather to characterize or describe the pedagogical approach that is evident through a MOOC's design.

We detail each dimension of the original AMP instrument:

- Epistemology. The epistemology dimension describes the extent to which instruction follows an objectivist approach or constructivist approach to instruction (Jonassen, 1991). Courses designed following an objectivist approach present instructional materials that are oriented toward acquiring knowledge and meeting course goals, while courses designed following a constructivist approach aim to create a rich learning environment where learners have opportunities to generate knowledge through social interaction with peers and draw on external resources in addition to predefined instructional materials.
- Role of teacher. The role of teacher dimension discerns the extent to which a course is teacher centered or student centered. In a teacher-centered course (Yuen & Hau, 2006), the majority of the instruction would be characterized as didactic. The instructor has a central role in the presentation of content and does not generally promote experiential learning or learner-to-learner interaction. Conversely, in a student-centered course (Zeki & Güneyli, 2014), the learning environment is designed to allow learners to showcase knowledge gained in multiple ways, through robust discussion and through peer-to-peer interaction facilitated by the instructor.
- Focus of activities. The focus of activities dimension sheds light on the process learners take to arrive at a solution to problems within the course. An activity is viewed as convergent if all answers are either right or wrong and there are no alternatives; otherwise it is divergent.

- **Structure.** The structure dimension describes the level of consistency within and across course units and subunits (i.e., weeks and lessons). It also evaluates the clarity of wayfinding information within a course. Highly structured courses are characterized by predictable course components patterns (e.g., quantity and sequence of components), and clear directions and navigational elements. This information provides learners with details about upcoming content and tasks, providing them with an indication of what to expect next.
- Approach to content. The approach to content dimension examines the extent to which course content is presented in a concrete or an abstract way. The focus of this dimension is not on whether the content itself is concrete or abstract but on how the instructor chooses to deliver and explain the content. In an abstract content presentation, material is presented as if it is self-explanatory. Concrete content presentation involves providing real-world examples to demonstrate a topic's application to everyday life, thereby making connections to learners' everyday lives.
- Feedback. The feedback dimension describes both the *frequency* with which learners receive feedback (frequently or infrequently) and the *types* of feedback learners receive (constructive or unclear). In the context of a MOOC, learners can have the opportunity to receive multiple forms of feedback within a course, both from the instructor and from their peers. When feedback is frequent, multiple opportunities for practice and graded assessments exist within a course. Constructive feedback can take different forms, including (1) elaboration on multiple-choice answers, which is written in advance by the instructor, and (2) personalized comments, given by other learners in the context of peer-graded assignments. Both of these types of feedback can supply information about whether an answer is correct or incorrect. They also offer additional information that learners can use to address and improve specific aspects of their assignment submissions.
- **Cooperative learning.** The cooperative learning dimension describes the extent to which cooperative learning (Roschelle & Teasley, 1995) is unsupported or integral within the design of a course. In a MOOC, cooperative learning can be instantiated in various course activities, including group projects, discussion board activities, and other forms of collaborative work among learners.
- Accommodation of individual difference. The accommodation of individual differences dimension describes the extent to which the design of a MOOC accommodates a wide range of individual differences, such as physical or cognitive disabilities. In a multifaceted course where accommodation of individual difference is supported, learners are able to present answers or outcomes through multiple means of representation, thereby directing their own learning (Burgstahler & Cory, 2010; Rose & Meyer, 2002). Course content can be accessed through more than one format (e.g., video, audio, or text).
- Activities/assessments. The activities/assessments dimension focuses on the characteristics of tasks within a course, specifically the context of the tasks: artificial or authentic. Tasks situated in artificial contexts usually have no strong connection to real life and in most cases only require declarative knowledge, such as the memorization of certain formulas and definitions (Wiggins & McTighe, 2005). Tasks situated in authentic contexts usually involve higher levels of Bloom's taxonomy (Anderson & Krathwohl, 2001), such as *create, apply*, and *evaluate*. By engaging in these tasks, learners have the opportunity to

reflect on the connection between the task itself and its application in real life (Herrington, Reeves, & Oliver, 2006).

• User role. The user role dimension discerns whether the course design promotes a passive or generative role for the learner. If the learner's role is passive, they primarily access content developed by the instructor. If the learner's role is generative, they may provide additional examples or links to external materials, enriching content developed by the instructor.

The AMP tool has been taken up by researchers that seek to characterize the pedagogies of individual MOOCs (e.g., Fan, 2017; Skrypnyk, de Vries, & Hennis, 2015). Skrypnyk et al. (2015) used the AMP tool to assess pedagogical dimensions of five MOOCs and found that they all differed in their learning design approach, such as the degree to which the designs were instructor or learner centered. In another example, Fan (2017) used the AMP instrument to assess pedagogical characteristics of five STEM and five non-STEM MOOCs from XuetangX and found differences between the learning design approach of STEM versus non-STEM courses. Swan et al. (2014) used the AMP tool and showed that MOOC pedagogies vary according to the course delivery platform, with courses hosted on Coursera being more instructor centered than courses on the Udacity platform. These MOOCs replicated traditional forms of instruction, including a lecture-based delivery mode followed by traditional assessments (e.g., multiple-choice quizzes).

Researchers have also made significant strides in grouping MOOCs that are similar, toward developing typologies. Swan et al. (2015) used the AMP instrument to derive three MOOC types from their analysis of 17 MOOCs: (1) participation oriented (MOOCs that provide many opportunities for learner interaction with content and other learners), (2) acquisition oriented (MOOCs that focus primarily on content delivery and automated assessment), and (3) self-directed (MOOCs that allow learners to determine their own pathway through a course). At the structural level, Davis et al. (2018) explored methods for clustering MOOCs that exhibit similar sequences of course elements (e.g., videos, readings, discussion prompts). They sought to use quantitative methods, such as transition probabilities and trajectory mining, to measure differences between course designs, thereby allowing course designs to be classified in a scalable fashion.

Machine-learning clustering methods hold promise for systematically grouping pedagogically similar courses, although this application is currently not commonly observed within MOOCs. Clustering with machine learning has the advantage of being able to find quantitative, mathematical relationships between data and grouping these data based on these relationships without any biases external to the mathematical data itself (Jain, Murty, & Flynn, 1999). Through machine learning, researchers can often expedite their data analysis or find novel methods of perceiving sets of data based on pure mathematical relationships (e.g., Luo, Wang, & Zhang, 2003). An example of clustering in MOOC research can be found in Hicks, Roy, Shah, Douglas, Bermel, Diefes-Dux, and Madhavan's (2016) analysis regarding characteristics of fully engaged learners in a MOOC. They first used machine-learning clustering methods to group learners based on numerical clickstream data. Learners were divided in terms of click frequency, and those found by the algorithm to click the most in significantly distinct clusters were defined by the researchers as fully engaged. Through their use of machine learning, the researchers were able to group learners based only on quantitative behavioral data. In the present study, we used machine learning to group MOOCs based on the numerical data generated across 10 pedagogical dimensions as defined by the AMP tool (rather than learner data). This approach allowed us to create groups that were based on pedagogical similarities across courses without any other influences (e.g., researchers' knowledge of course goals, subject matter).

Materials and Methods

Our study was conducted in three phases (see Figure 1):

- Phase 1: After review of AMP literature, elaborate and/or modify the AMP instrument
- Phase 2: Review of 20 MOOCs to test the efficacy of the Expanded AMP instrument
- **Phase 3:** Cluster analysis of 20 MOOCs, using scores from the Expanded AMP instrument as data inputs



Figure 1. Flowchart depicting the three phases of study.

Phase 1

We used descriptions of the AMP tool (Swan et al., 2014; Swan et al., 2015; Swan et al., 2016) as a foundation for this work, taking careful note of all available details. As we have described, the AMP instrument consists of 10 dimensions, with a scale of 5 for each dimension. Each dimension is anchored by two poles (e.g., abstract and concrete in the approach to content dimension)—see Table 1. Five score levels (i.e., 1–5) relate to the *focus* of each dimension, rather than providing a numeric ranking. For instance, in the epistemology dimension, a score of 1 relates to a highly objectivist approach, and a score of 5 relates to a highly constructivist approach, with neither approach being ranked higher than the other.

Table 1

Pedagogy dimension	Left pole	Right pole
Epistemology	Objectivist	Constructivist
Role of teacher	Teacher centered	Student centered
Focus of activities	Convergent	Divergent
Structure	Less structure	More structure
Approach to content	Concrete	Abstract
Feedback	Infrequent, unclear	Frequent, constructive
Cooperative learning	Unsupported	Integral
Accommodation of individual difference	Unsupported	Multifaceted
Activities/assessments	Artificial	Authentic
User role	Passive	Generative

Dimension Names, Poles, and Ordering of Dimensions From Swan et al.'s (2015) Original AMP Instrument

Two reviewers conducted preliminary testing of the original AMP instrument on five MOOCs (not included in this study) and achieved poor consistency in scoring. Reviewers compared written rationales and discovered that there were discrepancies in their interpretations of the AMP instrument. Through an iterative process, these reviewers fine-tuned the instrument, making deletions, modifications, and elaborations, resulting in the Expanded AMP instrument.

Phase 2

For this study, we chose a sample of 20 MOOCs across a variety of subject areas, including data science (six), social science (three), science (one), health care (three), education (three), computer science (two), and business (two). The selected MOOCs were hosted on Coursera (Coursera, n.d.) and were developed at the authors' home institution over a 4-year time frame. We wanted to include a range of learning designs in our sample. Thus, we selected courses that were developed over a 4-year time frame to represent the course platform feature changes over time. We also selected courses from seven different subject domains to capture the potential differences in learning design caused by content domains.

Two reviewers independently coded three MOOCs from our sample using the Expanded AMP instrument. Initial IRR was 60%. Through discussion, reviewers resolved differences and updated the instrument. Reviewers coded an additional 30% of the sample and achieved IRR of 90%. The first reviewer coded the remaining 55% of the sample, including recoding the initial three test cases. At the end of Phase 2, all 20 MOOCs were coded. See Figure 2 for an example of the application of the AMP instrument to a data science MOOC.

Pedagogical dimension	Score			lev	el		
Structure	Unstructured	1	2	3	4	5	Structured
Approach to Content Presentation	Abstract	1	2	<	4	5	Concrete
Characteristics of Tasks	Artificial	1	2	3		5	Authentic
Feedback	Infrequent, unclear	1	2	1	4	5	Frequent, constructive
Characteristics of Evidence	Convergent	1	2		4	5	Divergent
Epistemology	Objectivist	1	1	3	4	5	Constructivist
Instructor Role	Instructor-centered	1	L	3	4	5	Learner-centered
Learner Role	Passive	1	2	3		5	Generative
Cooperative Learning	Unsupported	×	2	3	4	5	Integral
Accommodation of Individual Difference	Unsupported	1	2	3	4	5	Multi-faceted

Figure 2. Scored pedagogy dimensions of a data science MOOC using the Expanded AMP instrument.

Phase 3

In order to group MOOCs assessed with the AMP instrument, we chose to use machine learning to investigate what kinds of relations exist among the MOOCs that we analyzed. Through using machine-learning methods, we wanted to generate MOOC groups using scores from the 10 dimensions defined by the AMP tool without any influence from qualitative information about the course. We used MATLAB R2018a, a numerical computing environment (MATLAB, 2018) to run our analysis. We applied nearest neighbor clustering methods (Philippe, Cojocaru-Mirédin, Duguay, & Blavette, 2010) to calculate the squared Euclidean distance (Dokmanic, Parhizkar, Ranieri, & Vetterli, 2015) between one course's scores on each dimension and every other course to locate the closest neighbors (i.e., courses that were most similar to each other). The squared Euclidean distance allowed the visual patterns in score levels, such as the one captured in Figure 2, to be represented in a quantifiable statistical measure, which further enabled the reviewers to investigate the pedagogical relationships between the 20 MOOCs as scored using the Expanded AMP instrument. With the value of the squared Euclidean distance on the 10 dimensions between the 20 courses, the nearest neighbor clustering methods allowed us to conduct an exhaustive search for the most similar courses and produced clusters based on the value of the squared Euclidean distance. The smaller the squared Euclidean distance was, the more similar the courses were. We performed a preliminary inductive analysis (Thomas, 2006) of the clusters toward understanding what *features* make groups of MOOCs similar (i.e., related to pedagogical design decisions) and what *factors* might contribute to the similarity of the groups (i.e., related to practical considerations). To perform this analysis, we identified prominent features of each cluster relating to pedagogical design decisions. We listed characteristics of courses within these clusters that set them apart from courses in other clusters, such as a focus on providing multiple opportunities for social interaction, an abundance of authentic examples and dramatizations, or a focus on asking learners to present declarative knowledge. To identify potential underlying factors that might influence the design of courses within a cluster, we listed key differentiating features relating to production time frame, team composition, or position of course within a series (if applicable).

Results

Expanded AMP Instrument

We made several substantial adjustments to the original AMP instrument to improve usability, including (1) standardizing and fine-tuning language; (2) modifying descriptions of score levels; (3) adding or elaborating on descriptions of score levels; (4) grouping related dimensions, thus changing the order of the dimensions from the original AMP instrument; (5) writing summaries of each dimension; (6) adding question prompts; and (7) adding a list of course elements that reviewers should focus on for each dimension. See Table 2 for details of these changes.

Table 2

Details and Examples of Our Modification/Elaboration of the Expanded AMP Instrument

Modification/ elaboration action	Description of change	Example(s)
1. Standardizing and fine-tuning language	Standardized and fine-tuned language of dimension names and pole names	We (a) flipped the pole names <i>Abstract</i> and <i>Concrete</i> to keep the focus of each side of the instrument consistent; we (b) changed the name of the <i>Approach to Content</i> dimension to <i>Approach to Content Presentation</i> so reviewers would deliberate about presentation approach rather than domain characteristics.
2. Modifying descriptions of score levels	Clarified focus of each dimension where (a) there appeared to be overlapping descriptions with other dimensions, (b) where we expected to see no variation from course to course, (c) and where reviewers were asked to "count" criteria to arrive at score level	We (a) removed "sequenced instruction" from the <i>Epistemology</i> dimension because it seemed to relate to the <i>Structure</i> dimension; we (b) removed "self-paced" from the <i>Role of Teacher</i> , because all Coursera MOOCs are self-paced; we (c) removed "count" criteria such as "2 of 4 criteria met."
3. Adding or elaborating criteria of score levels	Added or elaborated criteria of score levels (a) where they were missing and (b) where further clarification or explanation was needed	We (a) added the "frequency of opportunities for learners to get feedback" as an additional criterion in the <i>Feedback</i> dimension; we (b) further clarified "robust discussion is encouraged" by adding "if instructors have provided learners with specific directives to interact with other learners" in the <i>Instructor Role</i> dimension.
4. Grouping related dimensions	Reordered sequence of dimensions to create a more logical and intuitive flow and grouped related dimensions	We grouped Characteristics of Tasks with Feedback, Characteristics of Evidence with Epistemology, and Instructor Role with Learner Role.

5. Writing summaries of each dimension	Wrote summaries to introduce the goal of reviewing each dimension, to provide reviewers with specific goals and focus	For the <i>Feedback</i> dimension, we added, "This dimension asks reviewers to assess the characteristics and usefulness of the feedback provided (i.e., from instructor, peers, and platform) and the potential that it has to help learners improve performance on future tasks."
6. Adding question prompts	Added guiding question prompts that relate to the poles of each dimension	For the <i>Epistemology</i> dimension, we added, "Are instructional materials oriented towards acquiring knowledge? Are course goals predefined and absolute?" (objectivist approach) and "Is the instructional environment oriented towards generating knowledge through social interaction with peers? Is the learner voice evident through the construction of course goals?" (constructivist approach).
7. Adding a list of course elements that reviewers should focus on for each dimension	Added a list of course elements (e.g., quizzes, discussion prompts) that specifically relate to evidence required to score each dimension	For the <i>Characteristics of Evidence</i> dimension, we added "Quizzes" and "Peer-graded assignments."

We did not add or remove pedagogical dimensions or alter the number of score levels, which remained at 10 and five, respectively. See Figure 3 for an annotated view of the instructor dimension of the Expanded AMP tool. For each dimension, we identify four additions:

- 1. **Guiding question prompts related to the poles of each dimension.** These questions can help reviewers confirm whether a course is closer to one poll or the other for each dimension. The questions are structured such that they have "yes" or "no" answers. If the answer to a question is "yes," then this is a good indication that the course exemplifies qualities related to that pole. It is then up to the reviewer to decide how strong the evidence is in order to make a precise determination about the score level.
- 2. Criteria of each level in each dimension. Each score level includes a description of the characteristics that must be present in a course for it to be scored at that level.
- 3. A summary statement to introduce the goal of reviewing this dimension. This statement situates the dimension within the broader AMP instrument.
- 4. **Suggested course elements to be considered to find evidence of scoring this dimension.** These elements are not meant to be an exhaustive list, but they can direct reviewers where to look first within the course when scoring a particular dimension.

A complete set of all 10 dimensions of the Expanded AMP tool can be found in Appendix A.



Figure 3. Annotated view of Instructor Role dimension (1 of 10) of the Expanded AMP tool.

In order to provide reviewers with guidance on how to approach scoring each dimension, we elaborated on the evidence that they should consider for each dimension. Note that we present our new order here, which allows for a more logical flow, with related dimensions grouped together:

- Structure
 - Evidence for a structured course includes a similar length and learner workload across units; similar patterns of course components, sequence, and quantity across units; a clear and structured syllabus with detailed information, such as learning outcomes; and sufficient directions for learners to access resources or engage in activities. Evidence for an unstructured course includes imbalanced length and workload across units; no obvious patterns of course components, sequence, and quantity; and lack of wayfinding information.
- Approach to content presentation
 - Evidence for abstract content presentation includes content being presented as selfexplanatory without contextualizing concepts with examples. Instructors who adopt a concrete approach to present content usually introduce new concepts together with examples and scenarios.
- Characteristics of tasks
 - When tasks are situated in artificial contexts, learners are expected to only activate declarative knowledge without making connections to the real world. When tasks are situated within authentic contexts, the task setting usually involves real-world problems, and tasks outcomes can potentially be applied to learners' own work or life.
- Feedback
 - To review the *frequency* of feedback, reviewers should examine both the total number of opportunities to receive feedback (e.g., total assessments in a course) and also the distribution of these opportunities (e.g., the interval period between two assessments).

To review the *types* of feedback, reviewers should consider whether the feedback provides information that allows learners to self-correct or further improve.

- Characteristics of evidence
 - Evidence for convergent approaches would be that the majority of assessments are in the form of single-correct-answer questions, such as multiple-choice questions. Evidence for divergent approaches includes open-ended questions that can be answered in multiple ways or projects that can be approached from different perspectives.
- Epistemology
 - Evidence for objectivist approaches to instruction in a MOOC includes (1) instructional design that foregrounds direct instruction, with learners acquiring knowledge primarily through lecture-based modes of content delivery and (2) assessments that lack flexibility and activities do not allow for learners to make connections from personal experiences and to external resources. Evidence for constructivist approaches to instruction in a MOOC includes a rich learning environment that includes a variety of course elements (e.g., videos, discussion prompts) that are employed to deliver information and facilitate interaction and a flexible learning environment that enables learners to integrate personal goals and experiences into activities and assessments.
- Instructor role
 - Evidence for an instructor-centered MOOC includes the following: the majority of assessments are auto-graded with little or no human response; learners have no choice in how they represent assessed learning outcomes; and no specific directives are provided for learners to interact with peers. Evidence for a student-centered course includes the following: the majority of assessments are generative, and learners have choice in ways to demonstrate achievements; multiple robust discussion opportunities are provided and facilitated by instructors or provided by instructors through written directions prepared in advance.
- Learner role
 - Evidence for a generative learner role includes opportunities for learners to generate content and be inspired by the connections between course materials and external resources. Evidence for a passive role includes the lack of opportunities for learners to create content to share within a course. An example of a course design that promotes a passive learner role is one in which the majority of course content consists of lecture videos or readings and learners' contributions are not encouraged.
- Cooperative learning
 - To decide whether cooperative learning is supported in a course, reviewers should first look for evidence that collaboration opportunities such as group projects exist in a course; the next step is to examine whether detailed information and strategies are provided by the instructor to help learners be successful in those cooperative learning settings.
- Accommodation of individual difference
 - Evidence for accommodation of individual difference includes the use of multiple representations to deliver content, such as both text and video being used to describe an assignment, opportunities for learners to make choices in terms of presenting

answers, the flexibility to support self-directed learning, and the compliance of accessibility best practices, such as providing video captions and transcripts.

Larger MOOC Comparison

We used the scores generated from Phase 2 as data inputs for our nearest neighbor analysis. Figure 4 portrays the review of 20 MOOCs using the Expanded AMP instrument; this figure also underscores our need for a systematic method for evaluating pedagogical similarities, since visual comparisons alone would be difficult. Our nearest neighbor analysis of 20 MOOCs resulted in seven groupings of pedagogically similar MOOCs (Figure 5), with each group consisting of two to four MOOCs.



Figure 4. Top: scores of 20 MOOCs coded using Expanded AMP instrument. Bottom: key of working names for 20 MOOCs in sample.



Figure 5. Top left: squared Euclidean distance of 20 MOOCs represented in a 3-D plot, with nearest neighbors connected by red lines. Top right: top view of the 3-D plot to show the seven clusters of pedagogically similar courses. Bottom left: key of working names for 20 MOOCs in sample. Bottom right: seven clusters of pedagogically similar courses.

Our inductive analysis of these clusters led to the following *preliminary insights* related to pedagogical design decisions (features) and practical considerations (factors) that could account for the relationships we identified through the cluster analysis.

Features. Five of the seven clusters consisted of MOOCs from different domains (e.g., Cluster 5 contained a data science course and a social science course), while the remaining two clusters consisted of MOOCs from related domains (e.g., Cluster 1 contained one medical education course and two courses about higher education).

All clusters consisted of courses that were related along several pedagogical dimensions from the Expanded AMP instrument. Through our inductive analysis, we articulated three distinctive attributes that likely influenced design decisions:

- **consistency in course composition** (i.e., nearly identical sequences of elements, from week to week),
- **use of a theoretical perspective** (e.g., use of constructivist-oriented pedagogies, with many opportunities for learners to generate knowledge), and
- **implementation of signature pedagogies** (e.g., use of teaching methods that are known to prepare learners to be practitioners in a professional field, and imparting a set of beliefs about professional attitudes, values, and dispositions). (Shulman, 2005)

- **Factors.** We also considered factors that could influence pedagogical design in our inductive analysis (i.e., those not related to the dimensions that are reviewed using the AMP instrument):
- **courses created by same instructional design team** (i.e., learning designers who supported faculty members),
- **courses produced within the same time frame** (e.g., one cluster was produced before elaborative feedback for quizzes was available on the MOOC platform), and
- **courses created as part of a single specialization** (i.e., courses created by different instructors but bundled as a related group of courses).

Discussion

This study responds to Reich's (2015) challenge to MOOC researchers to move beyond studies of engagement (e.g., completion rates) and to move toward studies of factors that lead to *learning*. Using Swan et al.'s (2015) AMP instrument as a foundation, we developed the Expanded AMP instrument, which will allow researchers and learning designers to consistently characterize the pedagogies of MOOCs within their own contexts. We anticipate that this expanded tool will support learning designers who desire a vocabulary to talk about the style of instruction in a MOOC. Such a vocabulary can facilitate communication between learning designers and other stakeholders by providing points of reference. Learning designers can use this vocabulary as a guideline for the pedagogical design of a course as it is being built and also use it for the evaluation of MOOCs by giving a means to measure how the final course design aligns with original pedagogical goals.

The systematic and replicable approach that we present for grouping pedagogically similar courses addresses Reich's (2015) second challenge, that of progressing from studies of individual courses to making comparisons across multiple courses. Using the AMP tool along with clustering methods has made it possible for us to group pedagogically similar courses in a systematic and accurate manner. Our preliminary inductive analysis gave us a good starting point for understanding why courses may have been grouped together and has enabled us to understand the pedagogical features and potential underlying factors that likely contributed to the similarity of courses within each cluster. We are also working on refining our approach so that we can know more precisely how each pedagogical dimension contributes to the similarity of courses within a cluster, by identifying the score variation on each dimension.

Use of these tools and methods could lead to the development of additional MOOC typologies (i.e., building on Sfard's [1998] acquisition and participation models) that are based on a nuanced understanding of underlying pedagogies. These approaches could also support learning designers and researchers who want to explore program-level considerations. For example, as MOOC series are becoming more prominent, learning designers must attend to pedagogical considerations at both the course and the series level. We hope that our work creates a pathway from research to practice, where practitioners who are deeply involved in the design and evaluation of MOOCs can use these developing tools and methods to characterize the pedagogical approaches that underlie design, allowing them to tell powerful stories that are based on evidence.

References

- Admiraal, W., Huisman, B., & van de Ven, M. (2014). Self- and peer assessment in massive open online courses. *International Journal of Higher Education*, *3*(3), 119–128.
- Anderson, L. W., & Krathwohl, D. R. (2001). A taxonomy for learning, teaching, and assessing (abridged ed.). Boston, MA: Allyn and Bacon.
- Burgstahler, S. E., & Cory, R. C. (Eds.). (2010). Universal design in higher education: From principles to practice. Cambridge, MA: Harvard Education.
- Coursera: Take the world's best courses, online. (n.d.). Retrieved from http://www.coursera.org
- Davis, D., Seaton, D., Hauff, C., & Houben, G. (2018). Toward large-scale learning design: Categorizing course designs in service of supporting learning outcomes. In *Proceedings* of the Fifth (2018) Annual ACM Conference on Learning at Scale (L@S). ACM.
- Dokmanic, I., Parhizkar, R., Ranieri, J., & Vetterli, M. (2015). Euclidean distance matrices: Essential theory, algorithms, and applications. *IEEE Signal Processing Magazine*, 32(6), 12–30. doi:10.1109/msp.2015.2398954
- Fan, Y. (2017). Use the AMP tool to characterize pedagogical approaches taken by MOOC courses in mainland China. *International Journal*, 11(1), 141–146.
- Head, K. J. (2017). *Disrupt this! MOOCs and the promises of technology*. Lebanon, NH: University Press of New England.
- Herrington, J., Reeves, T. C., & Oliver, R. (2006). Authentic tasks online: A synergy among learner, task, and technology. *Distance Education*, *27*(2), 233–247.
- Hicks, N. M., Roy, D., Shah, S., Douglas, K. A., Bermel, P., Diefes-Dux, H. A., & Madhavan, K. (2016, October). Integrating analytics and surveys to understand fully engaged learners in a highly-technical STEM MOOC. In 2016 IEEE Frontiers in Education Conference (FIE) (pp. 1–9). IEEE.
- Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. ACM Computing Surveys (CSUR), 31(3), 264–323.
- Jonassen, D. H. (1991). Objectivism versus constructivism: Do we need a new philosophical paradigm? *Educational Technology Research and Development*, 39(3), 5–14.
- Luo, M., Wang, L. N., & Zhang, H. G. (2003). An unsupervised clustering-based intrusion detection method. *Acta Electronica Sinica*, *31*(11), 1713–1716.
- Major, C. H., & Blackmon, S. J. (2016). Massive open online courses: Variations on a new instructional form. *New Directions for Institutional Research*, 2015(167), 11–25.
- MATLAB R2018a (2018). The MathWorks, Inc., Natick, Massachusetts, United States.
- Najafi, H., Rolheiser, C., Harrison, L., & Håklev, S. (2015). University of Toronto instructors' experiences with developing MOOCs. *The International Review of Research in Open and Distributed Learning*, *16*(3).
- Philippe, T., Cojocaru-Mirédin, O., Duguay, S., & Blavette, D. (2010). Clustering and nearest neighbour distances in atom probe tomography: The influence of the interfaces. *Journal of Microscopy*, 239(1), 72–77.

- Quintana, R. M., Tan, Y., & Korf, N. (2018). *Visualizing course structure: Using course composition diagrams to reflect on design*. Paper presentation at the Annual Meeting of the American Educational Research Association. April 13–17. New York, New York.
- Reeves, T. (1996). Evaluating what really matters in computer-based education. In M. Wild & D. Kirkpatrick (Eds.), *Computer education: New perspectives* (pp. 219–246). Mathematics, Science and Technology Education Centre, Edith Cowan University, Perth.
- Reich, J. (2015). Rebooting MOOC research. Science, 347(6217), 34-35.
- Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In C. E. O'Malley (Ed.), *Computer-supported collaborative learning* (pp. 69–197). Berlin: Springer-Verlag.
- Rose, D. H., & Meyer, A. (2002). *Teaching every student in the digital age: Universal design for learning*. Alexandria, VA: ASCD.
- Seaton, D. (2016, January 29). Exploring course structure at HarvardX: A New Year's resolution for MOOC research [blog post]. Retrieved from <u>https://vpal.harvard.edu/blog/exploring-course-structure-harvardx-new-year's-Resolution-mooc-research</u>
- Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher*, 27(4), 4–13.
- Shulman, L. S. (2005). Signature pedagogies in the professions. Daedelus, 134(3), 52-59.
- Skrypnyk, O., de Vries, P., & Hennis, T. (2015). Reconsidering retention in MOOCs: The relevance of formal assessment and pedagogy. In *Proceedings of the Third (2015) European MOOCs Stakeholder Summit* (pp. 166–172).
- Swan, K., Day, S., & Bogle, L. (2016). Metaphors for learning & MOOC pedagogies. In Proceedings of Third (2016) ACM Conference on Learning at Scale (L@S) (pp. 125– 128). ACM.
- Swan, K., Day, S., Bogle, L., & van Prooyen, T. (2014). AMP: A tool for characterizing the pedagogical approaches of MOOCs. *E-Mentor*, *2*(54), 75–85.
- Swan, K., Day, S., Bogle, L., & van Prooyen, T. (2015). AMP: A tool for characterizing the pedagogical approaches of MOOCs. In C. J. Bonk, M. M. Lee, T. C. Reeves, & T. H. Reynolds (Eds.), *MOOCs and open education around the world* (pp. 105–118). New York, NY: Routledge.
- Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, 27(2), 237–246.
- Wiggins, G. P., & McTighe, J. (2005). Understanding by design. Alexandria, VA: ASCD.
- Yuen, K. M., & Hau, K. T. (2006). Constructivist teaching and teacher-centred teaching: A comparison of students' learning in a university course. *Innovations in Education and Teaching International*, 43(3), 279–290.
- Zeki, C. P., & Güneyli, A. (2014). Student teachers' perceptions about their experiences in a student centered course. *South African Journal of Education*, 34(3).

Appendix A

Expanded AMP instrument full version (consists of one scoring sheet and 10 scoring guide sheets with detailed explanation for 10 pedagogical dimensions respectively)

Pedagogical	category			Rati	ngs				
Structure			Unstructured	1	2	3	4	5	Structured
Approach to 0	Content Presentation	n	Abstract	1	2	3	4	5	Concrete
Characteristic	s of Tasks		Artificial	1	2	3	4	5	Authentic
Feedback			Infrequent, unclear	1	2	3	4	5	Frequent, constructive
Characteristic	cteristics of Evidence Convergent		1	2	3	4	5	Divergent	
Epistemology		Objectivist	1	2	3	4	5	Constructivist	
Instructor Rol	structor Role Instructo		Instructor-centered	1	2	3	4	5	Learner-centered
Learner Role			Passive	1	2	3	4	5	Generative
Cooperative L	earning		Unsupported	1	2	3	4	5	Integral
Accomodatior	n of Individual Differe	ence	Unsupported	1	2	3	4	5	Multi-faceted
5	Structure		Level 1		Le	evels 2	2-4		1/11 Level 5
• Ev	Structure idence (unstructur	red): Ar	Level 1	but any	Le	evels 2	2-4	interna	1/11 Level 5
• Ev org co • Ev ex ex	Structure ridence (unstructur ganized without cons urse? ridence (structured perience of the cours pect?) Are clear dire	ed): Ar sistency): Are u se cons	Level 1 re units presented withor y across units? Are clea units organized consiste sistent from unit to unit? and navigational eleme	out any ar direc ently (e ' (e.g., nts pre	Le v appar ctions a e.g., by can le esent w	ent con and nav elemer arners <i>i</i> thin th	2-4 isistent i igationa nt type, make a e course	interna al elem sequer reasor e?	1/11 Level 5 Il structure? Is the course ents missing from the noce)? Is the learners nable prediction of what to

Note: CCD in the bottom left refers to the Course Composition Diagrams that the reviewers created for each course in the sample (n = 20) by following the method developed by Quintana, Tan, and Korf (2018). CCD is a type of interactive visualization used by learning designers and researchers to represent the sequence and pattern of course elements.

Presentation	Level 1	Levels 2-4	Level 5
 Evidence (abstract): I used in the real world of Evidence (concrete): or everyday life? 	s content presented as if it is so or everyday life? Is content presented with exam	elf-explanatory, without example	es of how it relates to or is to or is used in the real world
This dimension asks reviewers to examine the extent to which content is presented/applied in a concrete or abstract way.	Abstract: Presentation does not include examples of how it relates to or is used in the real world or everyday life (i.e., material is presented as if it is self-explanatory)	Middle Level 2 The presentation of the majority of content is abstract Level 3 A balance between abstract and concrete presentation of content Level 4 The presentation of the majority of content is concrete	Concrete: Presentation includes examples of how this subject relates to or is used in the real world or in everyday life
Videos 🔍 Readir	ngs		3/
Videos Q Readir Characteristics of Tasks	ngs Level 1	Levels 2-4	3/ Level 5
Videos Q Readir Characteristics of Tasks • Evidence (artificial): declarative knowledge • Evidence (authentic	Are tasks situated in contexts t e, formulas, rules, or definitions): Are tasks situated in authenti Artificial:	Levels 2-4 that are not authentic? Do activi c contexts? Do tasks regularly i Middle	Level 5 ties ask learners to state nvolve real world problems? Authentic: Tasks are situated in

Feedback	Level 1	Levels 2-4	Level 5
 Evidence (infreque feedback unhelpful Evidence (frequent feedback useful (e.g) 	nt, unclear): Does the learner h (e.g., no elaborative feedback p , constructive): Does the learn g., clear, constructive, and perso	nave infrequent opportunities to re rovided, feedback is not personal er have frequent opportunities to onalized)?	eceive feedback? Is the ized)? receive feedback? Is the
This dimension asks reviewers to assess the characteristics and usefulness of the feedbac provided (i.e., from instructor, peers, and platform) and the potentia that it has to help learners improve performance on future tasks.	Unclear: Learner has infrequent opportunities to receive feedback; feedback is not clear or constructive (neither correct answers no elaborative feedback are provided); feedback does not address specific aspect of learners' contributions	Middle Level 2 Feedback is mostly infrequent and unclear Level 3 Level 3 infrequent/constructive or Elements of both infrequent/constructive feedback are present ts feedback is mostly frequent and constructive and constructive	Constructive: Learner has frequent opportunities to receive feedback; clear and constructive (correct answ with elaborative feedback personalized (directly addresses learners' contributions)
Q Quizzes Q Pee Characteristics of Evidence	r-graded assignments C	(Discussion prompts)	5/11 Level 5
Q Quizzes Q Pee Characteristics of Evidence • Evidence (convergent): possible answer (e.g., a • Evidence (divergent): answers (e.g., as with a This dimension asks	r-graded assignments Level 1): When learners are in the process with questions in STEM courses When learners are in the process ctivities in the humanities)? Convergent:	(Discussion prompts) Levels 2-4 cess of arriving at a solution, do the ses)? ss of arriving at a solution, do the Middle	5/11 Level 5 ney consider only one y consider multiple Divergent:

Epistemology	Level 1	Levels 2-4	Level 5
 Evidence (objectivist goals predefined and a Evidence (constructive through social interaction) 	approach): Are instructional m absolute? vist approach): Is the instruction on with peers? Is the learner vo	aterials oriented towards acqu onal environment oriented towa vice evident through the constru	iring knowledge? Are course rds generating knowledge, uction of course goals?
This dimension asks the reviewer to discem the epistemological thrust of a MOOC This dimension relates to what is provided to the learner AND what the learner does.	Objectivist: Focus on instructional materials that are oriented toward acquiring knowledge; and course goals are predefined and absolute	Middle Level 2 To large extent objectivist Level 3 Both objectivist and constructivist elements are present Level 4 To large extent constructivist	Constructivist: Focus is on creating a rich "learning environment" where learners draw from a variety of sources to acquir knowledge, including peers integration of learners' goals, experiences, and abilities into their learning environment
		0	۰
Videos Q Readir	ngs Q Discussion prom	npts Q Assignments (Levels 2-4	Syllabus 7/11 Level 5
Videos Readir Instructor Role • Evidence (instructor- characterized as a trac • Evidence (learner-cen- learners? This dimension relates to the role of the instructor, from instructor-centered	ngs Discussion prom Level 1 centered): Is the role of the ins litional, didactic approach. ntered): Is the role of the instruct Instructor-centered: No choice in ways learners represent knowledge and	Assignments Assignments Assignments Assignments Levels 2-4 tructor primarily to present con ctor primarily to facilitate the ex Middle Level 2 To large extent	Syllabus 7/11 Level 5 tent? This might be change of knowledge among Learner-centered: Choice in ways learners represent knowledge and

Learner Role	Level 1	Levels 2-4	Level 5
 Evidence (passive Evidence (genera course materials? 	e): Are learners required to acce tive): Are learners asked to gen	ess predetermined course cont erate content and make conne	ent (e.g., lectures and readings ections from external sources to
This dimension asks reviewers to evaluate th extent to which the learne role is passive or generative.	e The learner's role is primarily to access variou presentations from the instructor (e.g., lecture videos) and other course materials (e.g., readings)	Middle Level 2 s The learner's role is to a large extent passive Level 3 The learner's role is bot passive and generative Level 4 The learner's role is to a large extent generative	Generative: The learner's role is primarily to generate content, making connections from extern h sources to course materia
	nments Q Discussion pr	ompts	
CCDs Assig	nments Q Discussion pr	ompts Levels 2-4	9/1 Level 5
CCDs Assig	nments Q Discussion pr	ompts Levels 2-4 structures missing in the instru ures present in the instructiona	9/1 Level 5 Ictional design? al design?

Accomodation of Individual Difference	Level 1	Levels 2-4	Level 5
 Evidence (unsupport instructional design fai Evidence (multi-facet instructional design pro 	ed): Is application of universal c I to provide opportunities for lea ted): Is there evidence of both th ovide opportunities for learners t	lesign principles missing from t rners to make choices within th ne application of universal design to make choices within the cou	he course? Does the le course? gn principles? Does the rse?
This dimension asks reviewers to evaluate the extent to which the design of the MOOC accommodates a wide range of individual differences	Unsupported: universal design principles are not applied; instructions from the instructor are presented in <i>either</i> verbal or written formats; self-directed learning and individual choice in representation mode is not supported in the instructional design	Middle Level 2 Accomodation of individual difference is mostly unsupported Level 3 Some elements of a multi-faceted approach are evident Level 4 Accomodation of individual difference is to a large extent multi-faceted	Multi-faceted: universal design principles are applied (e.g., video is captioned); verbal and written presentations from the instructor are designed to work together (e.g., video and text describing an assignment); self-directed learning is allowed and supported; opportunities for learners to present answers through multiple modes of representation