welcome to the moocspace: a proposed theory and taxonomy for massive open online courses

Emily Schneider[1]

[1] Lytics Lab, Stanford University,
Stanford, CA
elfs@cs.stanford.edu

Abstract. This paper describes a theoretical framework and feature taxonomy for MOOCs, with the goal of developing a shared language for researchers and designers. The theoretical framework characterizes MOOC design goals in terms of stances towards knowledge, the learner, and assessment practices, taking as a starting point the affordances of the Web and digital learning environments. The taxonomy encompasses features, course structures, and audiences. It can be mapped onto the theoretical framework, used by researchers to identify similar courses for cross-course comparisons, and by instructional designers to guide design decisions in different dimensions. Both the theory and the taxonomy are intended in the spirit of proposal, to be refined based on feedback from MOOC researchers, designers, and technologists.

Keywords: taxonomy, knowledge organization, MOOCs, online learning theory

1 Introduction

If learning is the process of transforming external information into internal knowledge, the Internet offers us a universe of possibilities. In this context, MOOCs are simply a well-structured, expert-driven option for openly accessible learning opportunities. As of mid-2013, the boundaries of the moocspace[1] remain contested, with opinions (data-driven or no) generated daily in the blogosphere, the mainstream media, and an increasing number of academic publications. Meanwhile, decisions being made at a breakneck speed within academic institutions, governmental bodies, and private firms. What of the earlier forms of teaching and learning should we bring forward with us into networked, digital space, even as its interconnected and virtual

[1] Other types of open online learning opportunities that lend themselves to be named with similar wordplay include the DIYspace (e.g. Instructables, Ravelry, MAKE Magazine), the Q-and-Aspace (e.g. Quora, StackOverflow), the OERspace (indexed by such services as OERCommons and MERLOT), the coursespace (freely available course syllabi and instructional materials that are not officially declared or organized as OER), and the gamespace (where to even begin?). Then there is Wikipedia, the blogosphere and newsites, curated news pages (both crowdsourced, e.g. Slashdot, and personalized, e.g. Pinterest), and the great morass of affinity groups and individual, information-rich webpages.

nature allow us to develop new forms? How can an interdisciplinary, distributed group of researchers, course designers, administrators, technologists, and commentators make sense of our collective endeavor?

Towards a shared language for the *how* and *what* we are creating with MOOCs, I offer two frameworks. Firstly, for orientation towards the goals we have when we design MOOCs, I propose a theoretical framework that characterizes our assumptions about knowledge, the learner, and assessments. The framework takes as a starting point the affordances of the Web and digital learning environments, rather than those of brick-and-mortar learning environments.

Secondly, for grounding in the concrete, I offer a taxonomy of MOOC features, structures, and audiences, designed to capture the broad scope of MOOCs in terms of lifelong learning opportunities. Each element of the taxonomy can be mapped onto the theoretical framework to make explicit the epistemological stances of designers. The taxonomy can be used by researchers as a way of identifying similar courses for cross-course comparisons, and by instructional designers as a set of guideposts for potential design decisions in different dimensions. Finally, in the closing section of the paper, I provide an example of mapping the theory onto features from the taxonomy and introduce an application of the taxonomy as the organizing ontology for a digital repository of research on MOOCs, also referred to as the moocspace. Each framework is meant as a proposal to be iterated upon by the community.

2 A Proposed Theory *(Orientation)*

MOOC criticism and design decisions have largely been focused on comparisons with brick-and-mortar classrooms: how do we translate the familiar into these novel digital settings? Can classroom talk be replicated? What about the adjustments to teaching made by good instructors in response to the needs of the class? It is imperative to reflect on what we value in in-person learning environments and work to maintain the nature of these interactions. But to properly leverage the networked, digital environment to create optimal learning opportunities for MOOC participants, we also need to compare the virtual to the virtual and explore opportunities to embody the core principles of cyberspace in a structured learning environment.

Techno-utopian visions for the Web have three dominant themes: participatory culture, personalization, and collective intelligence. Participatory culture highlights the low cost of producing and sharing digital content, enabled by an increasing number of authoring, curatorial, and social networking tools [1]. In this account, personal expression, engagement, and a sense of community are available to any individual with interest and time—an ideal that MOOCs have begun to realize with well-facilitated discussion boards, and somewhat, with peer assessment. Some individual courses have also encouraged learners to post their own work in a portfolio style. But overall there are not many activities in this vein that have been formalized in the moocspace.

Participatory culture's elevation of the self is echoed in the personalized infrastructure of Web services from Google to Netflix, which increasingly seek to use recommendation engines to provide customized content to all users. The algorithmic

principles of this largely profit-driven personalization are extendable to learning environments, though desired outcomes for learning are more complex than the metrics used for business analytics--hence the need for learning analytics to develop robust and theory-driven learner models for adaptive environments. Visions of personalized digital learning include options for learners to engage with the same content at their own pace, or to be treated to differentiated instruction based on their preferences and goals [2]. In MOOCs this will require robust learner models based on interaction data and, likely, self-reported data as well. Analytics for this level of personalization in MOOCs have yet to be achieved but personalization is occurring even without adaptive algorithms, as distributed learners are primarily interfacing with content at their own machines, at their own pace. Finally, collective intelligence focuses on the vast informational network that is produced by and further enables the participatory, creative moments of the users of the Web [3]. Each individual learner in a MOOC enjoys a one-to-many style of communication that is enabled by discussion boards and other tools for peer-to-peer interaction. In the aggregate, this becomes many-to-many, a network of participants that can be tapped into or contributed to by any individual in order to share knowledge, give or get assistance with difficult problems, make sense of the expectations of faculty, or simply to experience and add to the social presence of the virtual experience.

These themes are embodied in a range of epistemological stances towards two core dimensions of learning environments: the location of knowledge and conceptions of the learner. Assessment is the third core dimension of the learning environment [4]. The technology enables a wide number of assessment types but the stances towards assessment follow not from the affordances of the Web but from the standard distinction between formative and summative assessments. However, instead of using this jargon, I choose language that reflects the nature of the interaction enabled by each type of assessment, as the central mechanism of learning in online settings are interactions among learners, resources, and instructors [5] Finally, it is important to note that this framework treats the instructor as a designer and an expert participant, which also leaves room for the expert role to be played by others such as teaching assistants.

> ***Knowledge*: Instructionist-participatory**
> Where are opportunities to acquire or generate knowledge? Does knowledge live purely with the instructor and other expert participants or does it live in the broad universe of participants? Who has the authority to create and deliver content? Is the learning experience created solely by the course designers or is it co-created by learners?
>
> ***Learner*: Personalized-Collectivist**
> Are learners cognitively and culturally unique beings, or members of a network? Do the learning opportunities in the course focus on the individual learner or on the interactions of the group?
>
> ***Assessment*: Evaluation-Feedback**

What opportunities are provided for learners to make explicit their progress in knowledge construction? Are assessments designed to tell learners if they're right or to give them guidance for improvement?

The poles of each stance, as named above, are opposed to each other epistemologically, but one end is not necessarily preferable to the other. The choice between each stance is predicated on what is valued by the designer in a learning environment or learning experience, and what is known about effective instruction and learning activities from the learning sciences. Each feature of the course can be characterized along one or more of these dimensions (see Section 4.1). This means that multiple stances can exist in the same course.

3 A Proposed Taxonomy *(Grounding)*

The proposed taxonomy includes two levels of descriptive metadata. The first level characterizes course as a whole and is meant to evoke the broad set of opportunities available for sharing knowledge with MOOCs. The second level takes in turn each element of the interactive learning environment (ILE) and develops a list of possible features for the implementation of these elements, based on current and potential MOOC designs. The features on this level can also serve as a set of guidelines of options for course designers. In multiple iterations of the course, many of these fields will stay the same but others will change. Most fields will be limited to one tag but others could allow multiple (e.g. target audience in General Structure).

The architecture and options for metadata on learning objects has been a subject in the field for quite some time, as repositories for learning objects and OER have become more common. While I am somewhat remiss to throw yet another taxonomy into the mix, I believe that it is important to represent the unique role of MOOCs in an evolving ecosystem of lifelong learning opportunities. Because the content and structure of a MOOC is not limited by traditional institutional exigencies of limited seats or approval of a departmental committee and accreditation agencies, it becomes a vessel for knowledge sharing, competency development, and peer connections across all domains, from computer science to music production and performance.[2] As a technology it is agnostic to how it is used, which means that it can be designed in any way that our epistemological stances guide us to imagine. Education has goals ranging from knowledge development to civic participation and MOOCs can be explicitly designed to meet any of these goals.

3.1 General MOOC Structure

On the highest level, each MOOC needs to be characterized in terms of its subject matter, audience, and use. Table 1 presents the proposed categories and subcategories for the General MOOC Structure. With an eye towards future interoperability, where

[2] That said, there is an ongoing conversation about integrating MOOCs back into the pre-existing educational institutions, so the taxonomy must ne conversant with these efforts while also representing the vagaries of the moocspace as a separate ecosystem.

possible I use the terminology from the Learning Resources Metadata Initiative (LRMI) specification [7], or note in parentheses which LRMI field the moocspace categories could map onto.

Table 1. Categories and Subcategories for General MOOC Structure

- Name (LRMI)
- Numeric ID (auto-generated)
- Author (LRMI)
  - Faculty member
- Publisher (LRMI)
  - Affiliated university or other institution
- Platform
- inLanguage (LRMI)
  - primary language of resource
- Domain (*about*)
  - Computational /STEM – CS, math, science, computational social sciences, etc.
  - Humanist – humanities, non-computational social sciences, etc.
  - Professional – business, medicine, law, etc.
  - Personal – health, thinking, speaking, writing, art, music, etc.
- Level (*typicalAgeRange* or *educationalRole)*
  - Pre-collegiate; basic skills (i.e. gatekeeper courses, college/career-ready); undergraduate; graduate; professional development; life skills
- Target audience (*educationalRole*)
  - Current students, current professionals, lifelong learners
- Use (*educationalUse* or *educationalEvent*)
  - Public course (date(s) offered), content for "wrapped" in-person course (location and date(s) offered)
- Pace
  - Cohort-based vs. self-paced (*learningResourceType* or *interactivityType*)
  - Expected workload for full course (total hours, hours/week) (*timeRequired*)
- Accreditation
  - Certificate available
  - Transfer credit

3.2 Elements of the Interactive Learning Environment (ILE)

The ILE is made up of a set of learning objects, socio-technical affordances, and instructional and community design decisions. These features are created by the course designers -- instructors and technologists – and interpreted by learners throughout their ongoing interaction with the learning objects in the course, as well as the other individuals who are participating in the course (as peers or instructors).[3] The features of the ILE can be sorted into four distinct categories: instruction, content, assessment, and community. Table 2 lists out the possible features of the ILE, based on the current trends in MOOC design. As stated, this is a descriptive list - based on

[3] The individual- and group-level learning experiences that take place in the ILE are enabled by the technological infrastructure of the MOOC platform and mediated by learner backgrounds (e.g. prior knowledge, self-regulation and study habits) and intentions for enrolling [8] as well as the context in which the MOOC is being used (e.g. in a "flipped" classroom, with an informal study group, etc.). The relationship of these psychological and contextual factors to learning experiences and outcomes is a rich, multifaceted research area, which I put aside here to foreground the ILE and systematically describe the dimensions along which it varies.

the current generation of MOOCs – but will be expanded in the future, both to reflect new trends in MOOC design and to take a normative stance on potential design choices that are based in principles of the learning sciences or interface design. Some of the features are mutually exclusive (i.e. lecture types) but others could occur simultaneously in the same MOOC (i.e. homework structure). Most features will need to be identified by spending some time exploring the course, ideally while it is taking place.

Table 2. Features of ILE

Instruction • Lecture ▪ "traditional": 1-3 hrs/wk, 20+ mins each ▪ "segmented": 1-3 hrs/wk, 5-20 mins each ▪ "minimal": <1 hr/wk • Readings • Simulations/inquiry environments/virtual labs • Instructor involvement – range from highly interactive to "just press play"	**Content** • Domain (in General Structure) • Modularized ▪ Within the course ▪ connected with other MOOCs/OER • Course pacing ▪ Self-paced ▪ Cohort-based
Assessment • In-video quizzes ▪ multiple choice vs. open-ended • Homework structure ▪ Multiple-choice ▪ Open-ended problems ▪ Performance assessments ▪ Writing assignments or programming assignments ▪ Videos, slides, multimedia artefacts • Group projects • Practice problems (non-credit bearing) ▪ Grading form–Quantitative, Qualitative • Grading structure (relevant to all credit-bearing assessments) ▪ Autograded ▪ Peer assessment, self-assessment, both ▪ Multiple submissions	**Community** • Discussion board • Social Media - Facebook group, Google+ community, twitter hashtag, reddit, LinkedIn, etc. • Blogs / student journals (inside or outside of platform) • Video chat (G+ hangout, Skype) • Text chat

4 The Taxonomy, Applied

4.1 Example of course mapping

Each course feature can be mapped onto one or more epistemological stances. The course overall can then be characterized by the overall epistemological tendencies of the course features. Table 3 provides an example.

Table 3. Mapping "Crash Course in Creativity" to the Taxonomy

General	Name: Crash Course in Creativity Author: Tina Seeling Publisher: Stanford Platform: NovoEd Domain: personal-thinking	Level: life skills Target audience: lifelong learners Use: public course (fall 2012) Pace: cohort-based - **collectivist** Certificate: yes
ILE and Stances		
Instruction	Lecture: minimal – 5-10 mins/wk to inspire group projects – **participatory** Readings: free, from her book - **instructionist**	
Content	Not modularized - **instructionist**	
Assessment	One individual creative projects – **participatory, individualist** Three group creative projects – **participatory, collectivist** Peer grading with qualitative comments–**participatory, feedback, collectivist**	
Community	Discussion board – **participatory, collectivist**	
OVERALL	**Participatory, collectivist, feedback**	

4.2 Stances to guide best practices and analytics.

The stances are not normative but do help specify which traditions of instructional and interface design should be turned to for guidance in best practices for designing resources. For example: instructionist lecture videos should follow the principles of multimedia learning, including balancing and integrating visual and verbal representations, relying on segmented (and learner-paced) narratives, and providing signaling mechanisms for the upcoming structure and content of a lecture. [9] The underlying epistemologies can also provide guidance about the type of analytics that are appropriate to for characterizing success in the design of the MOOC. For example, group-level outcomes may be more compelling for a collectivist MOOC – what is the overall level of interaction between learners, what kind of social networks form, with group projects can we characterize group composition or dynamics that lead to higher grades?

4.3 Centralizing distributed science: a short description of the moocspace

The taxonomy is a high-level, qualitative categorization of MOOCs that will allow for meaningful comparison across shared metrics about the courses. The taxonomy will be most usefully implemented in the *moocspace* – a digitized repository of knowledge about the research and production of massive open online courses – so named because it is an abstraction and reflection of the larger moocspace. The MOOC, abstracted, will be the central object of the moocspace, attached to standard metrics about the course, as well as reports on any research that has been done with data from that MOOC.[4] Variations in metrics could be related to aspect of the course

[4] Developing a small, meaningful set of shared metrics for MOOCs is currently an open question. Higher education in the US is characterized by enrollment rates at the beginning of

design, which are formalized in the taxonomy. Beyond descriptive data, a transparent, well-organized research base will enable an incremental and cumulative set of evidence from both exploratory studies (e.g. building learner models based on observational data) and experiments on the multiplicity of instructional and interface design features. A well-documented experiment in a small number of MOOCs could be replicated elsewhere by other researchers, and the findings could be synthesized by a third group by comparing results across variations in course features.

The moocspace could also be expanded to include the content of the MOOC itself, if licensing decisions are made that will allow MOOCs to become re-usable and re-mixable pieces of OER. This implementation would involve paradata on the uses of MOOC materials and incorporate a community aspect where faculty who use the materials could talk about what worked or didn't work in their courses. Finally, the MOOC object could also be attached to open datasets on MOOCs. The individuals who using such datasets may not be inside the academy, which underscores the need to build a structure for sharing newly developed knowledge back with the community.

If the moocspace is to be implemented, we will need develop consensus on the features in the taxonomy, as well as a strategy for tagging existing courses (crowdsourced? local experts?) and for adding new features to the taxonomy.

Acknowledgments. Much thanks to Kimberly Hayworth for her role in the initial brainstorming of MOOC features; thanks also to the Lytics Lab, Roy Pea, Zach Pardos, and the anonymous moocshop reviewers for feedback on earlier versions.

4 References

1. Jenkins, H. (2009) *Confronting the challenges of participatory culture: Media education for the 21st century*. Cambridge, MA: MIT Press.
2. National Educational Technology Plan. (2010). *Transforming American Education: Learning Powered by Technology.* Washington, DC: US Department of Education, Office of Educational Technology.
3. Lévy, P., & Bonomo, R. (1999). *Collective intelligence: Mankind's emerging world in cyberspace*. Perseus Publishing.
4. Brown, A. L., & Cocking, R. R. (2000). *How people learn*. J. D. Bransford (Ed.). Washington, DC: National Academy Press.
5. Anderson, T. "Towards a theory of online learning." (2004) *Theory and practice of online learning*: 3-31. Athabasca University, retrieved from http://cde.athabascau.ca/online_book/ch2.html
6. LRMI Specification Version 1.1 (Apr 28, 2013). www.lrmi.net/the-specification
7. Grover, S., Franz, P., Schneider, E. and Pea, R. (2013) "The MOOC as Distributed Intelligence: Dimensions of a Framework for the Design and Evaluation of MOOCs." In *Proceedings of the 10th International Conference on Computer Supported Collaborative Learning* (Madison, WI, June 16-19).
8. Mayer, R E., ed. (2005) *The Cambridge handbook of multimedia learning*. Cambridge University Press.

---

the semester, and persistence rates and completion rates over time. In addition to enrollment and activity rates initially and over time, for open courses it may be more appropriate to examine levels of engagement, time-on-task, or participation on the discussion forum.