Mentor Academy: Engaging Global Learners in the Creation of Data Science Problems for MOOCs

Rebecca M. Quintana, Christopher Brooks, Cinzia Smothers, Yuanru Tan, University of Michigan
Email: rebeccaq@umich.edu; brooksch@umich.edu; cinziavs@umich.edu; yuanru@umich.edu
Zheng Yao, Chinmay Kulkarni, Carnegie Mellon University
Email: zhengyao@cmu.gmail.com; chinmayk@andrew.cmu.edu;

Abstract: We engaged MOOC learners (mentors; n=120) as collaborators in the design and improvement of an online course through an activity called the Mentor Academy; an online instructional program where mentors create original problem sets for future iterations of a course. Following a design research methodology, we trace the development of the program's design and report on the features of a learning design that effectively support mentors in the problem creation process. Our analysis shows that mentors were offered three kinds of support: community building, logistical, and pedagogical. The following kinds of changes and additions were made to the design: expanded project requirements, improved technical infrastructure, and increased opportunities for interaction in discussion forums and live peer chat sessions. We present an analysis of problem sets for quality, diversity of subject area, and usability.

Introduction and Theoretical Foundations

Instructors and designers of Massive Open Online Courses (MOOCs) face a two-fold challenge in creating content for a *global* audience of learners: (1) it is time consuming to develop robust problems for learners who require multiple opportunities to practice emerging skills; (2) problems should be representative of learners who are situated in diverse contexts. The first challenge is a logistical and pedagogical one; the second challenge relates to the growing worry that MOOCs currently promote a version of cultural hegemony (Head, 2017).

To address both of these challenges we designed the Mentor Academy, a two-week instructional program that invited learners who successfully completed the University of Michigan's Introduction to Data Science in Python MOOC to be *mentors* and engage in the creation of problems for use by new learners in the next iteration of the MOOC. Our work builds on the concept of *learnersourcing* as a method of engaging with learners to create new learning experiences (Kim, 2015). Learnersourcing "is a form of crowdsourcing in which learners collectively contribute novel content for future learners, while engaging in meaningful learning experiences themselves" (Kim, 2015, p 3). We asked mentors to identify and then use subject matter and datasets that they felt would be compelling and relevant for learners who live and work in the same local context (i.e., country). Our program resulted in both a more *plentiful* and a more *diverse* set of problems than the instructional team would be able to provide on their own.

Our approach also shares common ideas with *communities of practice*, where members are mutually engaged in a joint enterprise, and work with a shared repertoire while actively negotiating the nature of the enterprise (Lave & Wenger, 1998). It also shares commonalities associated with *cognitive apprenticeship*, a model that adapts apprenticeship methods for the teaching and learning of cognitive skills (Collins, 2006). Tasks are situated in authentic contexts, and are "tightly coupled to the underlying competencies needed to carry out the tasks" (Collins, 2006, p. 54). In the Mentor Academy, mentors were apprentice instructors, acquiring the skills and competencies required to create problems suitable for use by novices in the domain of data science. These skills were modelled by the instructors of the program. The mentors' work was situated in an authentic context, that of an existing MOOC, and their contributions were seen as essential for the betterment of the MOOC. The following research questions guided our investigation:

1. What features of the learning design effectively support mentors in the creation of problem sets for use in the next-iteration of an online data science course?
2. What are the outcomes of the Mentor Academy program with respect to problem creation?

Methodology and Research Design

We recruited 120 mentors from the following countries: United States, India, Canada, China, Germany, and Brazil. We hosted Mentor Academy as four private courses on Coursera, one of the largest MOOC platforms. Instructors prepared video lectures that targeted pedagogical concepts such as constructivism, scaffolding, and writing effective problems. During each two-week session, approximately 30 mentors were asked to locate local data sets and write problems that leveraged those datasets. The problems were written in the Python programming language, using the Jupyter notebook environment. Mentors shared their problems on discussion boards and in synchronous video conferences with instructors and other mentors. We used an iterative design research methodology in which two-week sessions overlapped by one week to allow for rapid iteration from one design cycle to the next. Transcriptions of design debrief sessions with course instructors and the research team and design documentation (e.g., annotated wireframes, digital spreadsheets) were analyzed to understand (1) effective program supports and (2) improvement strategies. We analyzed the 39 problems that resulted from the four Mentor Academy sessions with respect to quality, diversity of subject area, and usability.

Results and Discussion

Features of learning design that were effective in supporting mentors

Our analysis revealed (a) three types of supports that were effective in supporting mentors and (b) five improvement strategies. Types of supports included: (i) community building (i.e., opportunities for mentors to interact with instructors and peers), (ii) logistical support (i.e., videos and documentation that explain the mechanics of the platform and organization of the program), and (iii) pedagogical support (i.e., instruction about pedagogy and feedback on problems). Improvement strategies included: (i) expanded project requirements, (ii) improved technical infrastructure, and (iii) increased opportunities for interaction with instructors in discussion forums and live peer chat sessions. Expanded project requirements and improved technical infrastructure led to increased diversity among problems; we allowed non-Wikipedia data sets to be used (e.g., local government data) and added technical capacity by allowing mentors to use URLs that contained accents. We gave mentors opportunities to share datasets and problems with instructors and peers in week one of the program, created additional opportunities for synchronous feedback, and provided mentors with improved scripts for video conferences with that included only mentors (not instructors).

Outcomes of the Mentor Academy program with respect to problem creation

Our analysis of problems with respect to *quality* revealed that generally problems stated a clear goal as well as details relating to product/performance that were required to solve the problem. However, problems often lacked supporting details and sufficient context that would motivate learners to solve the problem. Future versions of the Mentor Academy will consider how instructional resources can target this aspect of problem creation.

Our analysis of problems with respect to *diversity* uncovered four general topic areas: (1) society (e.g., immigration, population trends, child mortality reduction); (2) economics (e.g., government budget distribution, automobile industry); (3) individual in society (e.g., sports, healthcare, transportation); and (4) environment (e.g., air pollution, deforestation, and natural disasters). Some problems related to temporally relevant topics (e.g., hurricanes and California wildfires). We intend to perform a more nuanced analysis of these problems to understand the extent to which problems are relevant to learners from specific geographic regions.

Our analysis of problems with respect to *usability* uncovered four problem types, with respect to size:

- Micro (n=8): Small questions which could be completed quickly with one or two lines of code, suitable for use within an “in-video” quiz (i.e., embedded within a lecture, executable within the video player);
- Meso (n=11): The intended question size for the Mentor Academy program, which tended to be between 5-15 lines of code and practiced a single skill or topic;
- Macro (n=19): A series of questions which built on one another, suitable for use as an assignment;
- Tutorial (n=1): A form of worked example which provides a descriptive analysis, showing insights gleaned through data analysis in a narrative form, and which is intended to be read (and perhaps replicated) by the learner, but is not a problem to solve *per se*.

These problems will be integrated in the next iteration of the Introduction to Data Science in Python MOOC.

References

Collins, A. (2006). Cognitive apprenticeship. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 47-60). New York, NY: Cambridge University Press.

Head, K. (2015). The Single Canon: MOOCS and Academic Colonization. In *MOOCs and Open Education Around the World*, edited by Curtis J. Bonk, Mimi M. Lee, Thomas C. Reeves, and Thomas H. Reynolds, 190-202. New York: Routledge.

Kim, J. (2015). *Learnersourcing: improving learning with collective learner activity* (Doctoral dissertation, Massachusetts Institute of Technology).

Lave, J., & Wenger, E. (1998). Communities of practice: Learning, meaning, and identity. New York: Cambridge University Press.