

# Scaling Up Productive Disciplinary Engagement With Participatory Learning and Assessment

Daniel T. Hickey, Suraj Uttamchandani, and Joshua D. Quick,  
dthickey@indiana.edu, suddamch@indiana.edu, jdquick@indiana.edu  
Indiana University Learning Sciences Program

**Abstract:** This paper presents the results of a three-year program of design research that aims to scale highly interactive social learning in open online courses. This research started with simple features that had previously been refined in conventional online courses in order to foster more productive forms of disciplinary engagement. These features reflect five design principles for fostering productive forms of disciplinary engagement, motivating and assessing that engagement, and finally assessing the resulting individual understanding and achievement. This paper shows how these features were scaled up for use by hundreds of students in an open course and summarizes the impressive levels of disciplinary engagement and achievement that resulted. It also presents two new features that were introduced to allow self-paced participatory learning with little or no instructor involvement.

**Keywords:** personalized learning, learning analytics, assessment, analytic approaches

## Introduction

MOOCs and other open online courses have dramatically expanded opportunities to learn by significantly reducing the added incremental cost of additional students once the course is established. The so-called “xMOOCs” (“eXtended”) at edX, Udacity, Coursera, and elsewhere typically feature streaming videos, online readings, automated problem sets and quizzes, and peer discussion forums. The explosive growth of MOOCs led the *New York Times* to deem 2012 “the year of the MOOC” while also leading to new scrutiny. Studies showed that many were disappointed in the more limited interaction in xMOOCs (e.g., Khalil & Ebner, 2013), and one effort to include more interaction and group projects in Coursera was widely cited for going “laughably awry” (Oremus, 2013). Some observers had already commented on the difficulty of connecting with others in the “connectivist” “cMOOCs” that were designed specifically to support social interaction at scale (Mackness, Mak, & Williams, 2010). One study found that engagement in Coursera discussion forums declined significantly over time among completers due to information overload as discussion threads become unnavigable, and that instructor involvement actually worsened participation (Brinton et al., 2014). The lack of peer and social interaction was a prominent concern in the widely-cited “backlash” against MOOCs in 2013 (e.g., Kolowich, 2013).

MOOC proponents responded that the social experience in typical MOOCs was actually quite similar to what many students experience in the large lecture courses that are common at many college and universities. Indeed, the peer discussion forums available for many MOOCs are similar to the informal study sessions that many students organize themselves into for conventional lecture courses. Nonetheless, many MOOC organizations and instructors began investing significantly in improving the usefulness of their discussion forums by adding features, trained volunteers, and paid discussion leaders. By 2014, a significant progress was already underway towards more social interaction and peer learning in MOOCs and in learning at scale more generally. For example, a team at Stanford was developing a MOOC platform (now called NovoEd) which is explicitly based on social learning theory (Ronaghi, Saberi, & Trumbore, 2015). A major program of research at the Open University resulted in FutureLearn which supports “discussion-in-context” and “community-supported learning” (Parr, 2013), while a promising new strand of research supporting social interaction around peer assessment emerged (Kulkarni, Socher, Bernstein, & Klemmer, 2014). It is particularly noteworthy that some of these efforts to scale social learning are drawing insights from the Computer-Supported Collaborative Learning and Learning Sciences communities. This includes work at Carnegie-Mellon using intelligent conversation agents and social recommendation technology to enhance edX discussion forums (Rosé, Goldman, Sherer, & Resnick, 2015) and work at the University of Toronto to support inquiry-oriented project-based learning and user-contributed content in edX (Håklev, Slotta, & Najafi, 2015). Indeed, commentators have recently pointed to increased social learning as the “next challenge” in supporting learning at scale (e.g., Bryant, 2015; Parr, 2014).

## Participatory social learning and the Assessment BOOC

This paper is concerned with scaling a particular type of social learning. This type of learning builds on theories of situated cognition that emerged at the Institute for Research on Learning starting in the late 1980s. In particular, the social learning described in this paper builds on the “situative synthesis” advanced by Greeno and colleagues (Greeno, 1998). This perspective diverges from prior theories of cognition by focusing primarily on knowledge that resides in social and cultural practices (i.e., is “situated”) and then framing individual behavior and individual cognition as secondary “special cases” of this more socially defined knowledge. Building on Greeno’s notion of “engaged participation” in the communal construction of socially-defined knowledge, these perspectives value participation *as* learning and are thus often referred to as “participatory.”

The instructional framework that is the focus of this paper is currently referred to as Participatory Learning and Assessment (PLA). The PLA framework emerged from an extended program of design-based educational research using situative theories of cognition to uncover new ways of engaging learners, assessing learning, and evaluating programs. The PLA framework as it is presented here initially emerged in design studies of secondary language arts instruction (Hickey, McWilliams, & Honeyford, 2011) and educational videogames (Hickey & Jameson, 2013). Key features used to enact this framework online emerged in conventional graduate courses on *Assessment in Schools* and *Learning and Cognition* (Hickey & Rehak, 2013). This paper concerns a new series of design studies that explored how these features could be streamlined and automated to allow the same interactive social learning while making fewer demands on the instructor. This was intended to allow “big” open online courses with hundreds of students in the near term, working towards massive courses with thousands of students in the longer term.

A grant from Google was used to offer a “big open online course” (BOOC) in the fall of 2013 based on the first author’s existing *Assessment in Schools* course. The grant supported customization of Google’s *Coursebuilder* platform to scaling “wikifolios” and peer interaction features that had previously delivered in small courses using *Sakai*. As summarized in Hickey, Kelly, and Shen (2014), over 500 students (mostly practicing teachers and faculty) registered for the 12-week course and 160 completed the first assignment. Over 60 students (including eight for-credit students) completed all three modules (assessment *practices*, *principles*, and *policies*) to earn a web-enabled digital credential, a “badge” that could contain all completed work and interactions. The “Assessment BOOC” was again offered in the summers of 2014 and 2015. With the grant funding depleted, the open course was not as widely promoted and the efforts were primarily supported by the for-credit students completing the course as a part of their graduate degree. Most of the programming effort was directed at further automating course features to allow learners to experience interactive peer learning while placing minimal demands on the instructor. Additional features were added to allow self-paced learners. This ability to support self-paced learners while still offering them the interactive peer learning experience has been a major goal of this program of research. The weekly deadlines and relatively demanding assignments made it difficult for many open learners to keep up with the course. More generally, these features should eventually make it possible to provide entirely peer-to-peer experiences requiring instructor input, eventually leading to an entirely automated massive course featuring much (but likely not all) of the interaction describe below.

## PLA design principles and corresponding features

The PLA design principles coordinate activity across different kinds of interactions that afford different kinds of learning. Drawing on Hall and Rubin’s (1998) studies of mathematics, the principles distinguish between interactions that are *public* (visible to every participant), *local* (in public but between specific individuals), or *private* (between individuals). A fourth kind of interaction, *discreet* (i.e., unobtrusive), was added to highlight the core PLA assumption that achievement tests may be appropriate (and are indeed necessary to support some claims of competency), but should be used judiciously and presented inconspicuously.

These principles are all oriented towards supporting what Engle and Conant (2002) called *productive disciplinary engagement* (PDE). This framework presumes that engagement that is *disciplinary* involves both the declarative knowledge of the discipline (what experts *know*) as well as the social and cultural practices in which disciplinary experts engage (what experts *do*, in professional contexts). They further argued that disciplinary engagement that is *productive* generates numerous connections between that declarative knowledge and the learner’s experiences engaging in disciplinary practices. The first two PLA design principles directly follow from Engle and Conant’s design principles for fostering PDE. Central to this extension is the manner in which PLA uses engagement in disciplinary *practices* (which are more concrete and meaningful) to foster engagement with disciplinary *knowledge* (which is more abstract and challenging to learn). A central insight from the prior design research was that peers more readily question and discuss characterizations of disciplinary practices because they are generally not “factual” like disciplinary knowledge (Hickey & Rehak, 2013).

## PLA principle #1: Use public context to give meaning to knowledge tools

This first principle embodies the core situative assumption that the context in which disciplinary knowledge is learned and used is a fundamental part of that knowledge (Greeno, 1998). Students' own prior experiences, current interests, and future aspirations (i.e., their nascent disciplinary practices) are used to publically "problematize" the disciplinary knowledge of the course. This is consistent with Engle and Conant's first design principle: *problematize subject matter from the perspective of the learner* (Engle & Conant, 2002).

### Public wikifolios

All interaction is organized around public (to the class) "wikifolios." These are intact pages where each set of instructions can be hidden or revealed, and where participants respond to instructions in simple WYSIWYG editing boxes. The wikifolios engage participants in more open-ended disciplinary practices (which naturally foster discussion) rather than more specific declarative knowledge. The wikifolios generally use personalized disciplinary practices to engage students with the more specific disciplinary knowledge. In each wikifolio, disciplinary engagement is fostered with "ranking" features, which have proven to be a simple and scalable way of fostering social PDE. Students rearrange text boxes summarizing 3-5 elements of declarative knowledge that are elaborated in the textbook and/or open educational resources. Participants make sense of that information by ranking the relevance of each element relative to their curricular aim (discussed below) and/or professional role, and then justify that ranking. A particularly important aspect of this feature is that even when students lack the experience or understanding to rank something, they must engage with the knowledge to reach that conclusion. This prepares them to readily appreciate the rankings and rationales of peers with more experience. The expectation which has been repeatedly borne out in practice is that learners develop a routine whereby they do an initial ranking and rationale before looking at examples of peers who they interact with regularly.

Each wikifolio included other activities that reflected the textbook chapters. For example, each of the wikifolios in the "assessment practices" unit had students create example assessments using item writing guidelines before ranking the relevance of those guidelines. Other features had students simply summarize the big ideas of a text chapter or external resource. Each wikifolio also included a number of "optional" elements that the for-credit students were required to complete. This open format and the various features described next succeed in motivating participants to write a great deal. In 2013 the eight credential students wrote an average of 1398 words per wikifolio across their 11 wikifolios, while the 60 open learners who completed the course averaged 1207 words. Most impressive was that the 100 open learners who started but did not complete the course averaged 1137 words per wikifolio. In 2014, the average number of words per wikifolio doubled for the 12 credential students to 2820, while the averages for the 10 open completers (1377) and the 54 who dropped (1080) stayed roughly the same. The 2015 course maintained these levels of engagement for the 23 credential students (2374) and the 2 open completers (1783), while the 5 who dropped averaged 843 words per wikifolio.

### Self-contextualization at registration

When participants register for the Assessment BOOC, they are asked to select a primary educational *role* (teacher, administrator, etc.) and *setting* (secondary, college, etc.) and asked to define an initial *curricular aim* (a learning goal for a particular class they have taught or might teach) that embodies that role. The role and setting are used to automatically assign students to a networking group on the participant locator page described below. The curricular aim is automatically inserted into the first wikifolio assignment, which guides students through the process of using a text chapter and open educational resources to further refine that aim. This registration feature highlighted the personalized approach that the course would take, presumably discouraging registrants who were not serious or did not like that approach. Reflecting the core situative assumption that disciplinary practice and disciplinary knowledge reciprocally define each other, learners were instructed to continue refining their curricular aim in each wikifolio as their knowledge of classroom assessment expanded. Put differently, the refined curricular aim embodies each learner's growing understanding of their own professional practice; as their knowledge of assessment grows, their knowledge of relevant aspects of their professional practice grows as well. This growing understanding their own professional experience provides the context which grounds their learning of new and otherwise more abstract disciplinary concepts as the course progresses.

### Peer networking groups and peer location tools

A challenge for fostering peer learning at scale is helping participants find ideal peers to interact with. The participant locator display is a page that lists each participant along with hotlinks to their published wikifolios. Depending on what kinds of peers one is looking for, participants can be displayed according to networking groups, primary role, recent updates, and number of peer promotions. In 2014, the third wikifolio assignment was modified to invite students to extend their usernames and thus project additional identities that were not included

in the original networking groups (e.g., *librarian, unemployed math teacher*). More than half of the students now do so. In 2015, a new archiving feature let participants archive their wikifolios in a way that reflected their wishes for interacting with subsequent participants after they had completed the course. One setting displayed the archived work in green. This setting caused subsequent comments to trigger an email to the author and indicated the author *would* respond to comments. The yellow archive setting also caused subsequent comments to trigger an email but indicated that author *might* respond to comments. The red archive setting did not trigger email notification (but subsequent students were still free to endorse, promote, and comment. While students also have the option to delete their wikifolios when they complete the course, very few do so.

## PLA principle #2: Reward productive disciplinary engagement

This second principle assumes that productive forms of disciplinary engagement with resources, peers and instructors should be facilitated and rewarded. This principle uses situative theories of motivation and incentives (Hickey, 2003) to enact Engle and Conant's second and third PDE design principles: *give students authority over their disciplinary engagement* and *hold students accountable for their disciplinary engagement*. Engle and Conant's study, and the various subsequent studies that they and others carried out to further refine these principles, were conducted in conventional classrooms (where PDE could be modeled by the teachers and fostered in conversation). Although several features had been refined for enacting these two principles online in the prior courses, enacting them at scale in the BOOC required significant innovation. It is worth noting that students were instructed to engage in peer commenting, promotion, and endorsement, but there was no requirement or accountability of any sort beyond the collaborative reflection described below.

### Peer commenting

A key feature for supporting PDE online was having students and instructors interact with each other via threaded comments posted directly to wikifolios. This contrasts with discussion forums, which are relatively removed in time and space from the completed student work, and which routinely veer from the topic of the assignment and sometimes veer from the topic of the course. The scale of the BOOC and the introduction of self-paced learning called for a way for participants to readily locate (a) new comments on their own wikifolios, (b) replies to their comments on other participants' wikifolios, and (c) new course announcements and feedback. A hotlinked notification feature was added at the top of every page alerting each participant to all such developments and linking directly to the new activity.

When coupled with the peer locator tools, this feature made it quite simple for participants to efficiently engage with their peers and allowed the instructor to efficiently highlight (and therefore reward) exemplary wikifolios and discussion threads via the participatory feedback described below. In 2013, the for-credit students and open completers averaged 5.6 and 3.0 comments per wikifolio, respectively; in 2014, these two groups averaged 4.2 and 3.4 comments per wikifolio. In 2015, these two groups averaged 4.0 and 3.6 comments per wikifolio; across groups and years, comments averaged around 100 works. Coding of a representative subsample of comments from 2014 revealed that 92% of the comments referenced the topic of the chapter directly (Hickey, Quick, & Shen, 2015). While we lack a ready comparison, this seems like much more disciplinary engagement than is typical of many conventional online courses and most open courses.

### Peer endorsement and promotion

These features assume that conventional peer assessment is usually awkward and not particularly productive, because it focuses on "known answer" questions about disciplinary knowledge. While students dislike assessing whether peers "know," something, they do not seem to mind assessing whether peers "did" something. The peer endorsement feature allowed participants to endorse wikifolios as "complete." Peers simply clicked one of two buttons if the author had completed (a) all of the required elements or (b) all of the required elements plus the optional elements. The peer promotion feature allowed to students to promote one (and only one) wikifolio each week as being "exemplary." Peers were required to include a warrant for the promotion and the warrant was displayed alongside the name of the peer who awarded it. In 2013, 56% of the students who posted a wikifolio also promoted a peer; this proportion increased to 60% in 2014 and to 86% in 2015. These seem like very promising levels, particularly given that there was no accountability for either type of participation.

### Participatory feedback

A major challenge in scaling up wikifolios is that most useful examples and interactions are "buried" within individual wikifolios. In the 2013 and 2014 Assessment BOOCs, the cohorted format made it possible to have two types of announcements around the weekly wikifolio deadlines. The *early posts* announcements went up after the first 3-5 participants posted their wikifolios each week, invariably including some of the most experienced and ambitious participants. The instructor would reward the early posters with extensive comments, elaborations,

and pointed questions used to introduce more advanced content. The early posts announcement included hotlinks and encouraged other students to examine the early posts and comments, but only *after* getting started on their own wikifolio. The assumption is that the experience of starting the wikifolio would provide the other learners the personalized context that would help them make sense of the examples and additional content.

The *post-deadline* announcement went up a day or two after each weekly deadline. This highlighted and linked to the most widely promoted wikifolios and important discussion threads. The post-deadline announcement also included one of the most novel and potentially far-reaching features of the course. Each week the instructor and research assistants would examine the aggregated rankings by networking group for illuminating patterns. For example, when completing the *Validity* wikifolio in the *Principles* module, most of the educators concluded that *content-related* validity evidence was most relevant (because it concerned the content of an assessment), while most of the administrators concluded that *criterion-related* validity evidence was most relevant (because it concerned the relationship between scores and external criteria such as promotion). A bar graph displaying this information and hotlinks to examples helped participants review and revisit difficult concepts. For example, this explained that the few students who found *construct-related* validity evidence most relevant were doctoral students who were interested in constructs like self-efficacy. Comments on the announcements page and reflections confirmed that this feedback helped the educators and administrators more fully understand the (highly abstract) notion of psychological “constructs.” We believe that this innovation is particularly productive and a good example of how situative theories lead to useful features that help learners make sense of the most challenging elements of disciplinary knowledge.

While participatory feedback seems very promising, it was also one of the most laborious aspects of the Assessment BOOC. Even though the ranking by group data was downloaded to a spreadsheet, it was still necessary to locate, summarize, and link to good examples. In 2015, the two forms of feedback were combined into one weekly post to ease the workload. Comments on the announcement pages and interviews confirmed that many participants (a) found the feedback and examples interesting and useful, (b) were motivated by the possibility of getting mentioned, (c) wanted to see who got mentioned each week, and (d) used the feedback to help review for exams. Providing participatory feedback has proven challenging in moving to self-paced courses. Currently, self-paced learners are encouraged to examine current or archived examples of the existing early poster-feedback and the current wikifolios that received the most peer promotions. Current efforts will automatically display graphs of the rankings by groups in real time with links to promoted examples.

### Digital badges

One additional goal of this program of research was exploring the ways that open digital badges (evidence-rich web-enabled credentials) could motivate and recognize productive disciplinary engagement and achievement. Completing each of the three modules generated a badge in which learners could choose to include any and all of their work and interactions. Earning the three module badges generated the course badge (which contained the three module badges). These badges employed the new Open Badges Infrastructure metadata specifications. This meant that learners could readily share them over email or social media, which would display a hotlinked impact that would redirect the badge viewer to our course site and the archived work.

Each module badge can display the number of comments, endorsements, and promotions posted on each wikifolio, along with a link to the wikifolio as well as the content of those comments and promotions. In 2013-2015, one member of each networking group was awarded a *Leader* version of the badge for earning the most peer promotions. The Leader badges clearly served to motivate some students to post early and engage more deeply with their peers. Efforts are now underway to determine if and how leader badges might be awarded in a fully self-paced courses without cohorts of students in which to make this judgement.

### **PLA principle #3: Evaluate artifacts via local reflections**

Happily, the features used to enact the first two principles generate lengthy wikifolios and extensive peer commenting. Particularly regarding the credential students (and the corresponding expectations for accountability), this creates a new challenge of evaluating and grading all of these artifacts and interactions. The third PLA principle eschews any formal summative evaluation of the content of public artifacts and local interactions. This principle thus builds on existing assessment research that suggests “no marks” (i.e., ungraded) feedback (Harlen, 2007) and concerns about excessively detailed portfolio and performance assessment rubrics.

These prior assessment guidelines were reframed using sociocultural approaches to portfolio assessment (Batson, 2011; Habib & Wittek, 2007). This perspective leads to the assumptions that artifacts themselves primarily show what learners *did* (not what they *can* do in the future). This is because the many different routes to producing a given artifact means that it is very difficult to use artifacts themselves to support claims of proficiency. This principle also reflects the corresponding assumption that formal summative evaluation of

artifacts in order to make such claims undermines the formative goal of building individual and collective knowledge around the creation and discussion of those artifacts (Gitomer & Duschl, 1995). Thus, instead of diminishing learners' engagement with the disciplinary content by marking it as "right" or "wrong," wikifolios were graded through learners' reflections. The features used to enact this principle have been refined and studied extensively (Hickey, 2015). This is a crucial feature for portfolio assessment because it helps resolve the tensions between formative and summative functions described by Barrett (2010).

#### Contextual, collaborative, and consequential reflections

One of the optional wikifolio elements (required for credentials) consisted of three carefully worded reflection prompts for learners to answer once they had posted their wikifolio and interacted with peers. Building on Gresalfi, Barab, Siyahhan, and Christensen (2009), participants are instructed to reflect on their *contextual engagement* ("How suitable was your context for learning this knowledge?"), *collaborative engagement* ("Who else's work and whose comments helped you learn this new knowledge?"), and *consequential engagement* ("What will you do differently in your context and beyond as a consequence of learning this knowledge?").

The assumption here is that students who had not engaged productively and socially with the disciplinary knowledge of the course will be unable to draft a coherent and convincing reflection. In this way the reflections summatively assess one kind of learning (prior engagement) while formatively assessing another kind of learning (understanding the relationship between new disciplinary knowledge and a variety of disciplinary practices). The ultimate intention of these reflections is rooted in the anthropological notion of *prolepsis* (the way anticipated future events shape present activity; Cole, 1993); learners know they will need to reflect and so engage more deeply than they otherwise might. Put differently, we assume that because learners know that they will have to reflect on these nuanced – but important – aspects of engagement, they are motivated to think about these prompts in advance of the reflections as part of completing their own wikifolios.

This feature helped limit both time-intensive private feedback to students and extended review of completed student work. So long as the for-credit students posted their wikifolios on time and included a coherent reflection, they were given all of their points for the wikifolios (which counted for 55% of the grade). In practice, the process of privately awarding points to the for-credit students was nearly automated. Points were awarded and unless the student work was particularly weak, a boilerplate feedback statement was added.

#### **PLA principle #4: Let individuals assess their understanding privately**

PLA assumes that the more formal assessments that efficiently generate valid evidence of *prior* learning inevitably frame that knowledge in ways that limit the assessment's value for directly supporting *new* learning (Author, 2013a; 2015). This leads to a second assumption that public and local interactions should not take place around the more static representations of knowledge (i.e., known answer questions) in formal assessments, and thus that any formal assessment of knowledge should be carried out privately. A third assumption is that well-designed "curriculum-oriented" assessments are uniquely suited for letting participants figure out for themselves how much declarative knowledge that they have taken away from their prior engagement with the resources and peers. A fourth assumption is that such curriculum-oriented assessments that are of a reasonable length can only cover a fraction of the declarative knowledge that engaged learners should take away from each assignment. A fifth assumption is that such assessments can't really assess the extent to which students connected that knowledge to their own disciplinary practices, because such knowledge is so highly contextual.

#### Ungraded self-assessments

In 2014, ungraded quizzes featuring six to eight open-ended assessment items were added to each wikifolio. Students had to enter a response to each item in order to see the scoring key for the item. These formative assessments were entirely voluntary and students were encouraged to attempt the items from memory. The instructions recommended that students who were unable to answer more than one item from memory should re-engage with their classmates (starting with the public feedback) and the text before taking the module exam. These instructions have been repeatedly refined in an effort to maximize the formative benefit for engagement and to discourage students from memorizing the answers to those questions to prepare for the exam.

#### **PLA principle #5: Measure aggregated achievement discreetly**

This principle encourages using externally developed multiple-choice achievement test items for very specific purposes. It has proven to be one of the most controversial aspects of the PLA framework due to widespread concerns that such tests narrow curriculum and focus on shallow factual learning. The "distal" items are "standards oriented." The principle assumes that as long as the items are not "cherry picked" to tap into topics of the specific curriculum, they can be used to create an achievement test that is largely independent of the way a particular

course was designed. As such they are useful (and indeed necessary) for measuring learning within courses, comparing learning across different versions of the same course, and accurately documenting course improvement over time. By “discreet” this principle means unobtrusive and ephemeral; course assignments should never be directly aligned to achievement tests. In most cases students should only see their overall score. Most importantly, little if any course time should be devoted to instructing students on how to answer multiple-choice items. Such tests should feature items that go beyond factual knowledge and the items should be analyzed using the item analysis routines that are widely available.

#### Time-limited multiple-choice achievement tests

Each of the three modules in the Assessment BOOC included a timed multiple-choice exam consisting of items selected from multiple assessment textbooks’ item banks. The exams included many “best answer” (rather than “correct answer”) items which would be impossible to look up and figure out with the limited time available. Test takers only saw their score, and not the correct answer for each item. Item analysis was used to replace misbehaving items. Scores averaged around 85% with one or two perfect scores and a normal distribution.

Starting in 2015, participants who did not attempt the exam or did not attain a score of 80% only earned badges for *Assessment Practices*, *Assessment Principles*, and *Assessment Policies*; Participants who took each exam and attained at least 80% earned the *Expertise* version of the module badge and could choose to include that information in their badge. Participants who earned all three badges also earned the *Educational Assessment* Badge which contained the other three badges. To earn the *Educational Assessment Expertise Badge* participants had to earn two out of three expertise badges and attain a score of at least 80% on a comprehensive final exam. Credential students were required to complete all of the exams, and those scores counted towards 45% of their grade in the course (10% for each module exam and 15% for the final).

## Conclusions

The impressive levels of engagement and achievement with hundreds of learners in 2013 and with dozens of learners and minimal instructor involvement in 2014 and 2015 support several conclusions. Most generally, these findings suggest significant progress was made in scaling up participatory learning. While we are still not prepared for massive scale courses, we are unaware of any other effort to scale up open learning that resulted in these levels of disciplinary engagement, understanding, and achievement. Second, these findings support the conclusion that *some scaling should be done gradually*. In order to quickly scale up to massive numbers of users, most MOOCs and MOOC platforms were forced to sacrifice interaction and personalization; because the code behind them is already so complex, those platforms are now finding it challenging to incorporate new features to support social learning. Third, the steady pace of improvement over years supports the conclusion that *scaling should be done iteratively*. These efforts were directly shaped by newer design-based research methods that emphasize the development of “local” theories in the context of reform efforts. Furthermore we conclude that such iterative refinements should be done within a coherent theoretical framework. Our commitment to situativity allowed us to draw directly from other research in that tradition to generate useful insights and solutions. In particular we found the notion of PDE particularly helpful, because it let us evaluate our innovations in terms of their presumed or actual impact on disciplinarity and productivity of interactions.

A final point is that this approach seems likely to have a much more profound impact on learners’ professional identities than most existing scalable instructional models. We conclude with our conviction regarding the value of insistently connecting the learning of new disciplinary knowledge with a growing understanding of one’s disciplinary practices *and* the practices of one’s peers. We believe that these approaches represent an efficient and scalable way of achieving the “joint accomplishment of identity” as recently described by Hand and Gresalfi (2015). This conclusion in turn points to a major question going forward regarding the extent to which this approach can be used in other domains and contexts. It is currently being used successfully in English, social studies, and biology Courses at the fully online *Indiana University High School* (Itow & Hickey, 2015). Further, pilot studies in Secondary Algebra and Freshman Calculus have confirmed that the framework requires substantial revision for use in mathematics (Uttamchandani & Hickey, 2015).

## References

- Barrett, H. (2010). Balancing the two faces of ePortfolios. *International Journal of ePortfolio*, 3(1), 6–14.
- Batson, T. (2011). Situated learning: A theoretical frame to guide transformational change using electronic portfolio technology. *International Journal of ePortfolio*, 1(1), 107–114.
- Brinton, C. G., Chiang, M., Jain, S., Lam, H. K., Liu, Z., & Wong, F. M. F. (2014). Learning about social learning in MOOCs: From statistical analysis to generative model. *IEEE Transactions*, 7(4), 346-359.
- Bryant, T. (2015, June 22). Bringing the social back to MOOCs. *EDUCAUSE Review*. [Online]

- Cole, M. (1993). Remembering the future. In G. Harmon (Ed.) *Conceptions of the human mind. Essays in honor of George A. Miller* (pp. 247–265). New York: Psychology Press
- Engle, R. A., & Conant, F. R. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learner's classroom. *Cognition and Instruction*, 20(4), 399-483.
- Gitomer, D. H., & Duschl, R. A. (1995). Moving toward a portfolio culture in science education. *Learning science in the schools: Research reforming practice*, 299-326.
- Greeno, J. G., & Middle School Mathematics through Applications Project Group. (1998). The situativity of knowing, learning, and research. *American Psychologist*, 53(1), 5–26.
- Gresalfi, M., Barab, S., Siyahhan, S., & Christensen, T. (2009). Virtual worlds, conceptual understanding, and me: Designing for consequential engagement. *On the Horizon*, 17(1), 21-34.
- Håkløv, S., Slotta, J. & Najafi, H. (2015, Oct). *It Wouldn't Be the Same without You: MOOC Design that Supports User-Contributed Content, Interaction and Teamwork*. Columbia University: Learning with MOOCs II,
- Habib, L., & Wittek, L. (2007). The portfolio as artifact and actor. *Mind, Culture, and Activity*, 14(4), 266-282.
- Hall, R., & Rubin, A. (1998). There's five little notches in here: Dilemmas in teaching and learning the conventional structure of rate. *Thinking practices in mathematics and science learning*, 189-235.
- Hand, V., & Gresalfi, M. (2015). The joint accomplishment of identity. *Educational Psychologist*, 50, 190–203.
- Harlen, W. (2007). Criteria for evaluating systems for student assessment. *Studies in Educational Evaluation*, 33(1), 15-28.
- Hickey, D. T. (2003). Engaged participation versus marginal nonparticipation: A stridently sociocultural approach to achievement motivation. *The Elementary School Journal*, 401-429.
- Hickey, D. T. (2015). A situative response to the conundrum of formative assessment. *Assessment in Education: Principles, Policy & Practice*, 22(2), 202-223.
- Hickey, D. T. & Jameson, E. (2012). Designing for participation in immersive educational videogames. In D. Ifenthaler, D. Eseryel, X. Ge (Eds.), *Assessment in game-based learning: Foundations, innovations, and perspectives* (pp. 401-430). New York: Springer
- Hickey, D. T., Kelly, T. A., & Shen, X. (2014). *Small to big before massive: Scaling up participatory learning and assessment*. Proceedings of the Fourth International Conference on Learning Analytics and Knowledge, Indianapolis, IN (pp. 93-97) <http://dx.doi.org/10.1145/2567574.2567626>.
- Hickey, D. T., McWilliams, J. T., & Honeyford, M. A., (2011). Reading Moby-Dick in a participatory culture: Organizing assessment for engagement in a new media era. *Journal of Educational Computing Research*, 44 (4), 247-273
- Hickey, D. T., & Rehak, A. (2013). Wikifolios and participatory assessment for engagement, understanding, and achievement in online courses. *Journal of Educational Media and Hypermedia*, 22 (4), 229-263.
- Itow, R. C., & Hickey, D. T. (2015, April 20). *Impacting online teacher practice, epistemic frames, and agency with design-based implementation research*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Khalil, H., & Ebner, M. (2013). “How satisfied are you with your MOOC?” A research study on interaction in huge online courses. In *World Conference on Educational Multimedia, Hypermedia and Telecommunications* (Vol. 2013, pp. 830–839). Retrieved from <http://www.edlib.org/p/112057/>
- Kolowich, S. (2013, May 1). Faculty backlash grows against online partners. *Chronicle of Higher Education*.
- Kulkarni, C. E., Socher, R., Bernstein, M. S., & Klemmer, S. R. (2014, March). Scaling short-answer grading by combining peer assessment with algorithmic scoring. In *Proceedings of the first ACM conference on Learning@ scale conference* (pp. 99-108). ACM.
- Mackness, J., Mak, S., & Williams, R. (2010). The ideals and reality of participating in a MOOC. In *Proceedings of the 7th International Conference on Networked Learning 2010*, 266-275.
- Oremus, Will. (2013, February 3). Online class on how to teach online classes goes laughably awry. *Slate*.
- Parr, C. (2013, May 23). FutureLearn reveals big plans to deliver MOOCs on the move. *Times Higher Education*, 23, 9.
- Parr, C. (2014, November 13). Making MOOCs social is the next challenge. *Time Higher Education*.
- Ronaghi, F., Saberi, A., & Trumbore, A. (2014). NovoEd, a social learning environment. In Paul Kim (Ed.), *Massive Open Online Courses: The MOOC Revolution* (96-105). New York: Routledge.
- Rosé, C. P., Goldman, P., Zoltners Sherer, J., & Resnick, L. (2015). Supportive technologies for group discussion in MOOCs. *Current Issues in Emerging eLearning*, 2(1), 5-20.
- Uttamchandani, S. L., & Hickey, D. T. (2015, September 11). Calculus PLANet: Promising first steps in participatory supplemental instruction in mathematics. *Remediating Assessment Blog*: <http://remediatingassessment.blogspot.com/2015/09/calculus-planet-promising-first-step-in.html>