Adventuring into MOOC Writing Assessment: Challenges, Results, and Possibilities

This article shares our experience designing and deploying writing assessment in English Composition I: Achieving Expertise, the first-ever first-year writing Massive Open Online Course (MOOC). We argue that writing assessment can be effectively adapted to the MOOC environment, and that doing so reaffirms the importance of mixed-methods approaches to writing assessment and drives writing assessment toward a more individualized, learner-driven, and learner autonomous paradigm.

Composition faculty find themselves ancillary to the political reform movements promoting assessment within their own institutions. They blink an eye, and suddenly some victor is nailing a manifesto right on their classroom door, directly affecting their programs and the lives of their students. The only recourse seems frustrated acceptance or angry protest. But . . . [w]e discovered, perhaps mostly by luck, that the real site of writing assessment is not so much a battle zone or a contested economic sphere of influence as a territory open for venture, and that writing teachers will do well neither to accept passively nor to react angrily— but simply to act.

Richard Haswell and Susan Wyche-Smith, “Adventuring into Writing Assessment”

Copyright © 2016 by the National Council of Teachers of English. All rights reserved.
Massive Open Online Courses (MOOCs) have widespread reach: “Almost 20 million learners in over 203 countries have enrolled in a [MOOC]” (Karsenti). Proponents insist that MOOCs will usher in a new, welcome era for higher education, with increased access, lower costs, and more options for learners (Daniel; Weissmann). Critics raise concerns about profit-driven motives, elitism, the automation of education, and the lack of personal interaction between faculty and students (Davidson; Farin; Manjikian; Canavan). We also have concerns about pressures to award college credit to students completing MOOCs without faculty supervision or valid assessment.

The stakes surrounding MOOCs are especially high with introductory-level courses. Dave Cormier predicts, “we could easily have 1 million (or 100 million) students taking first year physics online with MIT. . . . [I]ntroductory courses are obvious targets for the x-style MOOCs. All we’re really looking for is a general understanding of a given topic, you could do the testing in a Pearson test centre, pay $350, bang you’ve got a first year credit.” First-year writing is particularly vulnerable, as it is often targeted during institutional and political shifts because so many composition faculty are contingent labor and because first-year writing has so frequently been the subject of highly visible, fractious debates over where writing fits within curricula (Edgington 107). Although compositionists have long been leaders in online education (“CCCC Position”; Hewett et al.), concerns about the impact of MOOCs on first-year writing include potential labor loss and challenges to the essential nature of writing instruction: MOOCs, because of their massiveness, video lectures, and lack of sustained, one-on-one faculty-student interaction, seem inherently antithetical to writing pedagogy and thereby pose a threat to the core principles by which we hope to work with postsecondary writers. Ongoing developments in Automated Essay Scoring (AES) make the prospect of assembly-line writing instruction ever more possible and alarming (Herrington and Moran; Hesse), signaling increased mechanization of teaching and assessment and eliding human interaction around writing. Four MOOCs, in fact, experimented in 2013 with AES built into the EdX platform (Reilly et al.), and recent research maintains AES “holds the potential to play a viable role in high-stakes writing assessments” (Shermis 53).

White’s first law of assessodynamics”: Assess thyself or assessment will be done unto thee.

Edward M. White, “The Misuse of Writing Assessment for Political Purposes”
While the initial hype over MOOCs has settled in some ways, MOOCs are persisting. Even though Udacity’s Sebastian Thrun shifted focus to corporate training after remarking that their MOOCs were a “lousy product” (qtd. in Usher), other markers of MOOCs suggest they are going strong. At the end of 2015, Coursera was offering 1,471 courses through 136 institutional partnerships (Coursera website).

MOOC growth resonates in many ways with circumstances described over twenty years ago by Richard Haswell, Susan Wyche-Smith, and Edward White. In such contexts, those involved with first-year writing can choose to sit passively by, respond with anger, or, more constructively, “pay attention” (Selfe) and “act” (Haswell and Wyche-Smith). We decided to act. When the Bill & Melinda Gates Foundation issued a Call for Proposals for the development of introductory-level MOOCs across all disciplines, Denise Comer saw the prospect of designing a first-year writing MOOC as an opportunity to cultivate conversations about writing among learners around the world and to conduct research: What is the impact of MOOCs on writing pedagogy? What is the impact of MOOCs on student learning outcomes?

These questions drove the development of what became the first-ever first-year writing MOOC: “English Composition I: Achieving Expertise,” which ran for twelve weeks (18 March – 10 June 2013). It was funded largely through the Bill & Melinda Gates Foundation and offered through Duke University in partnership with Coursera. Course enrollment began at 64,000+ and climbed to 82,820 by the final week. Of these, 1,289 people earned a Statement of Accomplishment, which required a final grade of at least 70%; 689 of these learners earned a Statement of Distinction, which required a grade of at least 85%. We are neither surprised nor alarmed by this rate of completion. Many MOOC learners do not intend to complete a course, instead approaching it as an “experience experiment” or “for skill-enhancement purposes or personal self-actualization” (Voss 4). Given our course’s rigor and length, we expected many enrollees never to attend or engage with the steady demands of the course (see Jordan; Kolowich, “Coursera,” for more on MOOC completion rates).

Demographic data indicate that most enrollees (79.57%) lived outside the United States, held a bachelor’s or master’s degree (66.32%), were aged 26–44 (53.31%), and were employed full-time (86.91%). Learners enrolled from over 187 countries. Some 77% indicated that English was not their first language; however, 59% of these respondents indicated they were proficient or fluent in written English. These data parallel other research on MOOC student demographics (Voss).
The course had a single primary faculty instructor, Comer, but included a team of more than twenty people who contributed expertise in such areas as online learning, English as a Second Language, MOOC technology, disciplinary-based writing, assessment, and library studies. Edward M. White led our assessment efforts, designing and implementing many of the in-course assessments for student writing and the post-course assessment measures and evaluation. (See Appendix A for the full team.) Comer conducted research with IRB-approved protocols where participants could opt in to the research.

An experience with tens of thousands of writers from across the world yields numerous exciting possibilities to learn more about who writes and why, how we can or cannot teach writing to widely diverse populations, and how writing is valued and deployed across a range of cultural, demographic, geographic, and disciplinary contexts. We have chosen here to focus on what we learned regarding writing assessment: How can we design effective, consistent, and reliable writing assessment in a structure with tens of thousands of learners, each with markedly different levels of writing preparation and expertise, and where the writing teacher’s role largely excludes direct interaction with students and their writing? How do we adapt writing assessment to the MOOC context and, in so doing, what can we learn about MOOCs, writing, and writing assessment?

We see writing assessment as a form of research (Huot) with the potential to “improve the educational experience for students [through] greater understanding among faculty about their goals and expectations for student writers” (Zawacki and Gentemann 49). In the context of MOOCs, White’s assertion about the importance of assessment has especially high significance: “assessment is too important and its implications too far-reaching to be left to assessors and other specialists in measurement” (Teaching 135). If we in rhetoric and composition do not take the lead assessing writing in MOOCs, others—likely with less expertise or investment in writing and higher stakes in competing agendas—will.

Yet, the disruption presented by MOOCs also invites us to reexamine writing assessment at its core. Such a reconsideration aligns with Jeff Rice’s “Networked Assessment,” in which he argues that “new media logic” “favors tracing over finality . . . because its focus is on shifts in connectivity (or lack of connectivity) rather than on conclusive moments that remain fixed” (32). Rice’s argument propels us to look not only for the value of writing in the MOOC, but for “a series of relationships” (36). The relationships that emerged throughout
our experience included those among diverse learners and between teacher and student, even as they also yielded insights into relationships that might be forged between MOOCs and more traditional educational settings for first-year writing instruction.

In this article, we share our experience designing and deploying writing assessment in the English Composition I: Achieving Expertise MOOC, present our assessment results, and discuss the conclusions we have drawn and the relationships we have traced from this research. We argue that, for all the purported disruptiveness of MOOCs, many aspects of writing assessment used in more traditional face-to-face, online, or blended educational settings can be effectively deployed in the MOOC context. However, MOOC writing assessment must also be adapted to the practical realities MOOCs present. Doing so reaffirms the importance of mixed-methods assessment and reshapes writing assessment to incorporate a much more learner-autonomous and individualized approach. MOOCs, ultimately, sponsor an assessment paradigm where learners themselves hold the authority to define, construct, dismantle, and redefine their own frameworks and parameters for assessment, learning, competency, success, and failure.

**Course Overview: English Composition I: Achieving Expertise (EC)**

EC (see Appendix B for syllabus⁴) provided an introduction to college-level writing. The course challenged conventional classifications of MOOCs, operating at the nexus between cMOOCs and xMOOCs, which differ along many domains, particularly in such areas as the role of instructor, student demographics, flexibility of course material, and reliance on learner interaction (Rodriguez). The course included some elements of xMOOCs: “offered on university-based platforms [and] modeled on traditional course materials, learning theories and higher education teaching methods” (Morrison). Yet, it also involved several defining features of cMOOCs: “based on the explicit principles of connectivism (autonomy, diversity, openness and interactivity) and on the activities of aggregation, remixing, repurposing and feeding forward the resources and learning” (Rodriguez).⁵ The course adapted writing pedagogy into a cMOOC context by structuring content “based on the idea that learning happens within a network, where learners use digital platforms . . . to make connections with
content, learning communities and other learners to create and construct knowledge” (Morrison).

Essentially, it was up to the learners themselves to define where their experiences lay on the xMOOC to cMOOC continuum. Such a fusion reflects a point recently made by Tharindu R. Liyanagunawardena, whose “experience in [a MOOC suggested] that a MOOC may become ‘just an OER’ [Open Educational Resource, i.e., xMOOC] for some students if they do not form ‘MOOC friends’, a learning network, or participate in the MOOC discussions.”

This fusion of MOOC approaches posed a disruption within an already disruptive context. Because EC needed to offer an xMOOC experience, even if individual learners chose to make it a cMOOC experience, we grounded learning objectives (Table 1) in those established within writing studies (“WPA Outcomes”; Framework).10

Throughout EC, we focused on the value of meaningful peer feedback and on the importance of the writing process. We provided over 77 instructional videos (most of which were 4–6 minutes in length) about aspects of writing corresponding to the course’s arc: developing a claim, integrating evidence, incorporating visual elements, responding to the work of others, revision strategies, etc. (see Appendix E for video titles). EC asked students to draft and revise four major projects, participate in discussion forums, post lower-stakes writing assignments, produce self-reflections, and provide substantive peer review. We also invited students to participate (by Google Hangouts) in optional live writing workshops with other students and a member of the instructional team.

EC asked students to approach writing expertise by examining expertise itself. For their major writing projects, students chose an area of expertise of interest to them (a hobby, sport, art, discipline, career, etc.). The driving ques-
tions of the course enabled students to better understand their chosen area of expertise by writing about it for others: What factors engender achievement? Who determines success? These inquiries, however, worked twofold by also sponsoring our collective inquiry into writing expertise: What does expert writing look like? How does it vary across contexts? How does one develop writing expertise? Who defines what is or is not effective writing?

The first three projects went through a formal drafting, feedback, and revision cycle, and the fourth had an informal drafting stage:11

- Project One: Critical Review of Daniel Coyle’s “The Sweet Spot” (600–800 words)
- Project Two: Image Analysis (600–800 words)
- Project Three: Case Study of Expertise (1000–1250 words)
- Project Four: Op-Ed (500–750 words)

(See Appendix F for details of the projects.)

**Challenges of Writing Assessment in a MOOC**

Because no direct assessment model for writing in MOOCs was available, we hypothesized that we could adapt available assessment models to the MOOC context, thus creating a hybrid form of writing assessment. There is a venerable tradition of effective writing assessment as established by such scholars as Richard Haswell, Peggy O’Neill and Brian Huot, Katherine Blake Yancey, and Edward White. We designed assessments in accordance with established recommendations (“Writing Assessment”). We also consulted strong models for effective writing assessment in online environments (Cargile Cook and Grant-Davie; Miller-Cochran and Rodrigo; McKee and DeVoss; Bourelle et al.). And, by focusing on the relationships among learners, teacher and learners, and different contexts for writing instruction, we also invoked Rice’s notion of networked assessment.

Still, many of these models assume a closed, for-credit learning context, and there are no models (of which we are aware) for the kind of scale involved in MOOCs. Chief among the differences generated by the MOOC context are a lack of direct teacher feedback and evaluation, and limited accountability for peer feedback. Moreover, the scale of the MOOC and the wide heterogeneity of learners in terms of culture, age, experience, and motivation required a unique blend of structure and flexibility. An open assignment, for example, does not
seem appropriate for a MOOC, where students seem hungry for guidance. But the structure must be flexible enough to accommodate a wide diversity of experience, interests, and special knowledge. More broadly, the MOOC disrupts two key foundations for writing assessment: the importance of what Brian Huot terms a “local context” and the value of having what Terry Myers Zawacki and Karen Gentemann stress as discipline-based writing assessment.

Ultimately, because the MOOC was providing writing instruction to learners in an xMOOC manner, but relying extensively on cMOOC pedagogy, we needed to build a new framework for assessment, a hybrid approach that put into practice effective measures of writing assessment focused on value and validity, even as it also opened space for a flexible, learner-driven, and exploratory approach to assessment. Our approach, in essence, became one grounded on individuated assessment, where we aimed to facilitate the space and structure wherein learners might adapt, define, challenge, and reconstitute for themselves their own approaches to writing and learning assessment.

**Assessment Design Overview**

We designed assessments to be consistent, aligned with learning objectives, and useful for writers. We used multiple assessment measures and engaged the students in “contextualized, meaningful writing” (“Writing Assessment”). Through deliberate sequencing, our assessments asked students to intentionally connect smaller learning objectives for each unit to larger course objectives. Assessments foregrounded self-reflection and peer feedback, and all led toward a capstone self-reflective essay to facilitate autonomy and progress toward learning objectives.

Assessment design included the following specific components, each described in more detail below along with results:

- Peer Feedback and Evaluation
- Self-reflections
- Pre- and Post-course Self-efficacy Surveys

Ultimately, because the MOOC was providing writing instruction to learners in an xMOOC manner, but relying extensively on cMOOC pedagogy, we needed to build a new framework for assessment, a hybrid approach that put into practice effective measures of writing assessment focused on value and validity, even as it also opened space for a flexible, learner-driven, and exploratory approach to assessment.
Study Limitations
Study results for pre-and post-course surveys are self-reports, which may affect their reliability. Our holistic portfolio rating samples consisted only of people who had completed the course and agreed to participate in the IRB study, which may have skewed results. Additionally, the portfolios, added after the course had begun, did not achieve as retrospective and holistic a presence in the course as if we had integrated them more thoroughly from the beginning.

Results and Lessons Learned
We have included in this section qualitative and quantitative results and lessons learned for each component of the assessment design identified above. The value of well-structured, effective peer feedback for writing has been demonstrated repeatedly (Liu and Carless; Falchikov; Trim; Topping; Flynn, “Students”; Savic). Similarly, peer feedback in online environments, when deployed with best practices, has been demonstrated to be highly effective (Formo and Neary; Mabrito; Jordan-Henley and Maid; Breuch and Racine).

Literature on peer feedback emphasizes that directive instruments are crucial in determining efficacy (Flynn, “Re-viewing”; Gielen et al.; van Zundert, Sluijsmans, and van Merrienboer). For each of the four major projects, we designed highly structured rubrics: “formative” feedback for drafts (feedback toward revision [Black and William]) and “evaluative” feedback for final versions. We supplemented peer-feedback rubrics by modeling effective response with direct instructor feedback on posted samples of student writing and instructor-facilitated writing workshops on Google Hangouts, which were recorded and posted. All peer-feedback rubrics began with a statement emphasizing the benefits of peer feedback for the responder: “Reading and Responding to Other Writers Makes You a Better Writer.”

Formative peer-feedback rubrics (Appendix G) asked students to identify and comment on specific features of the writing project in alignment with the project and course learning objectives; these also included open-ended questions whereby peers could address author queries and reflect on how the peer responding impacted their own writing.
Our evaluative peer-feedback rubrics (Appendix G) were designed to assess the writing project’s success toward meeting the learning objectives for that assignment. Adopting features of effective scale development as articulated by such scholars as Sara Cushing Weigle, evaluative peer-feedback rubrics featured a six-point scale and had open-ended questions as well as a formative component: “What overall comments do you have for the writer as he or she moves on to Project 4?”

For evaluative feedback, we could have used Calibrated Peer Review (CPR) (Carlson and Berry; Kulkarni et al.) or explored Automated Essay Scoring (AES), but neither was a good fit for EC. AES largely eliminates human readers, whereas we adopted the premise that what student writers need and desire above all else is a respectful reader who will attend to their writing with care and respond to it with understanding of its aims. While an expert response is desirable, it is not necessary. Writers want readers, the more the better; AES precludes this. With CPR, students examine feedback models and then provide feedback in a “calibration” phase until they align closely enough to the model and can proceed to the evaluation phase. CPR, we believed, would have overemphasized evaluative feedback at the expense of formative feedback and, on a practical level, would have been too time-intensive given that we had four writing projects.

Peer feedback worked as follows: When a student submitted a draft, three other students (each of whom had also submitted drafts) received that draft and provided feedback according to the formative feedback rubric. After one week, the original author received the three sets of feedback and had one week to revise, whereupon the final version was distributed to four other classmates (each of whom also submitted a final version). These four peers used an evaluation rubric to rate the paper according to a scale (1 = poor quality; 6 = excellent quality). The Coursera platform drops the lowest of the four grades (in case a student receives an unwarranted low grade) and averages together the other three to arrive at a project grade.

Table 2 shows that learners who submitted a writing project were likely to participate actively in peer review (essentially 90%), suggesting they were invested in giving and receiving peer feedback.

These quantitative data, however, do not attest to quality. Based on a random sampling and learner impressions of peer feedback, the quality was uneven. As one student wrote, “In a way, it was like tossing a coin: sometimes you got lucky with your reviewers, other times you weren’t.”

Some peers were adept at identifying writing moves and offering suggestions toward revision. In the following example, the writer was apparently
dealing with a racially sensitive issue, and the peer suggests expanding racial perspectives:

Where does the writer analyse the image? Is that sufficient to convey the important aspects of the image to readers who may not see the image or have time to examine it thoroughly?

You analyse the picture in the fourth, sixth, and eighth paragraphs. I think they are sufficient and the first paragraph to illustrate this picture, in particular, is effective for readers who might be unfamiliar with Brazilian culture to understand your analysis. However, I would like to see more analysis about what class the people who exploit laborers belong to, the social status of Afro-Brazilian descendants in the Brazilian communities, why African-Brazilian have settled down in Brazil. You also should analyze a white worker with an empty bucket who has the same ethnic feature as the black laborer you mentioned. If they are delivered, I suppose readers will see real aspects of the Brazilian society more from the picture.

Another set of feedback shows responders deploying feedback with respectful, constructive, and positive tones. Peer 2 even applies what he or she has read in the peer’s project to a real-life experience.

peer 1 → Cool opening!

peer 2 → I very much liked your title, your paragraphs are clearly structured, references are included. I like your point of the teacher being someone who observes and guide the children, that is very different to what I have experienced as a child where in the classroom, we either did what we were told or were punished (by getting low mark or note).

peer 3 → i think that you have a good project
I admire your style of writing. Lucid, concise, and well structured. Well done you.

These examples show that many peers were taking their work as reviewers seriously.

In turn, many learners appreciated peer feedback: “I have found that one of the most helpful parts has been the peer feedback. Would anyone be interested in a peer feedback group once the class is over?”; “I did not expect many lines of feedback so I am very delighted with my peers feedback 😊 thank you!” Students expressed that peer feedback helped them focus on the learning objectives, revise effectively, and consider ideas from multiple perspectives: “I can be in the shoes of my readers and know how they perceive my writing” (IRB-Research Participant); “after going through peer review processes, I think more about my purpose, my audience and the way I want to organise my work.”

Feedback, however, also emerged in less productive ways, with negative tones:

*What did you learn from reading this essay?*
peer 1 → Nothing.

Some feedback was contradictory or minimal:

*Did you find the introduction effective?*
peer 1 → yes, it create effective introduction
peer 2 → No, because it is not related to the development of ideas. For tips see above

*Did you find the conclusion effective?*
peer 1 → yes, it tries to summaries the concepts well
peer 2 → The same goes for the conclusion

Based on the 90% participation in peer feedback among those who submitted a draft, and the fact that each person’s draft purportedly had three readers, one would hypothesize that each writer would get at least some meaningful feedback. However, some students expressed discontent with feedback overall: “But for the peer feedback, [the course] was very well organized.”

Dissatisfaction with peer feedback often stemmed from perceptions (realistic or not) that reviewers had inferior skills or lacked adequate English proficiency: “peer evaluation can be very frustrating, especially when . . . peers are not able to evaluate your work mainly because English is not their first language and they can not understand what they read.”
Continued efforts to strengthen the accountability of and support for meaningful peer feedback in a MOOC are important. As one student wrote, “I have given peer feedback for 15 drafts and evaluated 20 final projects from anonymous peers. That was a really good way to learn about what works and what does not.”

Peer assessment on final versions suggested that student writing overall was successful. We were uncertain initially if peers would grade excessively high or low, but Table 3 shows the average grades were between 4 and 5, the upper half of a six-point scale, not actually too different from the grades our expert readers tended to give in the portfolio rating (see section below).15

Our distribution and frequency data (Figures 1–4) show that peers made use of the full range of the rubric scale.

<table>
<thead>
<tr>
<th>Average Score on Final Version</th>
<th>Project 1</th>
<th>Project 2</th>
<th>Project 3</th>
<th>Project 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1=Poor, 6=Excellent. See Appendix G for Sample Evaluative Rubric)</td>
<td>4.51</td>
<td>4.55</td>
<td>4.60</td>
<td>4.49</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>.86</td>
<td>.81</td>
<td>.78</td>
<td>.78</td>
</tr>
</tbody>
</table>

Figure 1. Project 1 Peer Grade Frequency
Figure 2. Project 2 Peer Grade Frequency

Figure 3. Project 3 Peer Grade Frequency
Self-reflection is vital to writers’ growth (Pajares; Walker), and we agree with claims that student perceptions matter in the assessment of writing (Edgington 108). We sought to integrate many occasions for meaningful self-reflection, modeled on best practices from writing-studies scholarship (Smith and Yancey). Specific self-reflection components included the following:

- Discussion Forums
- Notes to Readers upon Draft Submission
- Peer-feedback Reflection
- Post-project Reflection
- Final Reflective Essay

Self-reflection provided one of the richest growth areas.

**Discussion Forums**

In discussion forum posts, many learners expressed enthusiasm for and appreciation of writing, while also being candid about their struggles as writers: “I write in two languages. My teacher once said: ‘To speak two languages is to live two lives.’ As I write more in another language other than my native one, I
think more in a different way and see a different self of me. My life is enriched because I am living another life. Writing helps me practice and choose words better in both languages.” Learners seemed motivated to improve their writing: “I’ve always had the unrealistic notion that I’d be a great writer if I wasn’t too lazy to try and secretly scared to fail. Well it’s time to give up that delusion: I’m a bad writer, but I won’t wait anymore for some magical inspiration, I’ll just write.”

Thousands of posts demonstrate a keen awareness of writing processes: “I have many bad habits when I’m writing, especially while the process is not happy. When I fall into deep thought, I often grab my hair or bite my nail. It’s really bad because my hair become less and my hands are not clean all the time.” These posts then invited ongoing conversation between learners about these challenges.

While Google Hangouts workshops had limited space, we emphasized that watching a Hangout enabled viewers to apply what they learned to their own writing. Table 4 shows thousands of Hangout views, though far fewer (145) posted reflections on what they learned. This suggests that more work can be done to emphasize the importance of reflecting on Hangouts. Those who did reflect did so meaningfully: “The key takeaway for me [from the workshop] was the balance between summary and evaluation; this was a point of indecision in my submitted draft.” Another student writes, “I picked up [from the Hangout] significant tips that I’m intending to use in my next exercise of analyzing a visual image. . . . I learned how to become more critical and pay close attention to details . . . enriching me not only on an academic but also on a personal level.”

### Table 4. Google Hangouts Workshop Views

<table>
<thead>
<tr>
<th>Writing Workshop Hangout</th>
<th>Views</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large-Group, Comer, Week 3</td>
<td>7792</td>
</tr>
<tr>
<td>Large-Group, Vidra, Week 3</td>
<td>6666</td>
</tr>
<tr>
<td>Small-Group, Comer, Week 3</td>
<td>5786</td>
</tr>
<tr>
<td>Small-Group, Erol, Week 6</td>
<td>4625</td>
</tr>
<tr>
<td>One-on-One Tutoring, Undergraduate Writing Tutor, Week 6</td>
<td>4503</td>
</tr>
<tr>
<td>Small-Group, Vidra, Week 9</td>
<td>3779</td>
</tr>
<tr>
<td>Large-Group, Jarmul, Week 11</td>
<td>3117</td>
</tr>
<tr>
<td>Small-Group, Font-Navarrete Week 11</td>
<td>1987</td>
</tr>
</tbody>
</table>
“Notes to Readers” upon Draft Submission

As writers submitted project drafts, we asked them to write a note to their readers:

**Note to Readers:** What specific questions or concerns would you like your readers to address regarding your draft? Would you like them to look at a specific passage? Would you especially like feedback about a certain aspect of your writing or argument? Were you stuck with any portions of the draft that you would like feedback on?

Their notes demonstrate high investment in getting feedback and self-awareness identifying elements needing more attention: “Do you think that the arguments presented in this case study are enough to support my theory? What do you think about the quotations presented? Is the case clearly developed? When I want to use a quotation taken from a web site how [do I] reference it?”

**Peer-feedback Rubrics**

One question on the peer-feedback rubric asked, “What did you learn about your own writing from reading this writer's essay?” Table 5 shows a sample set of responses and reflects the ways in which learners identified gains from providing peer feedback:

Responses such as these suggest that many peer reviewers were productively thinking about themselves as writers. Such a perspective is especially important in a MOOC, where the skill levels, experience, and backgrounds vary so much across individuals and where the feedback people receive can be uneven.

<table>
<thead>
<tr>
<th>Table 5. Sample Self-reflection on Providing Peer Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>peer 1 → This project helped me to appreciate the complexity of writing and communicating effectively to readers. As a writer, I believe the number one goal is clarity. This includes short words, simple sentences, active voice verbs and correct grammar. This project also reminded me of how writers make compelling arguments, provide supporting claims and evidence, and connect the points or evidence throughout the project.</td>
</tr>
<tr>
<td>peer 2 → I will learn how to paraphrase others’ work better.</td>
</tr>
<tr>
<td>peer 3 → I learned that citations can add very different dimensions to writing.</td>
</tr>
<tr>
<td>peer 4 → I learned that I should be a little skeptical while reading any author’s work and do some independent research while reviewing his/her book.</td>
</tr>
</tbody>
</table>
Post-project Reflection
Table 6 shows that a majority (~70%) of students who submitted writing projects also completed post-project self-reflections.16 (Project 3 and 4 post-project reflections were integrated in the final reflective essay.)

These reflections show learners identifying areas of potential growth and thinking ahead to the next writing project: “For future writing and particularly for the next assignment, I would like to be able to organize my essay well and not stray from the objectives of the assignment”; “MLA still confusing at times. I still have problems accessing journal articles and other sources from online or brick-and-mortar libraries.”

Final Reflective Essay
The final reflective essay was the culminating project, demanding a high level of awareness of the course goals and metacognition to assess one’s own performance. A total of 1,413 students submitted a final reflective essay. The examples in Table 7 demonstrate that students reflected on the learning objectives and provided evidence from their own writing for how they demonstrated growth in or the meeting of a particular objective.

Pre- and Post-course Self-efficacy Surveys
We offered pre- and post-course self-efficacy surveys asking students to indicate their confidence with aspects of writing related to our course learning objectives. Literature on self-efficacy, informed largely by the social-cognitive theory of Albert Bandura, suggests that confidence levels are directly correlated to student engagement, resiliency, and growth (Pajares). Because EC was introductory level, we expected enrollees would have low- to mid-level confidence in themselves as writers. In fact, we found the opposite. As demonstrated in Table

<table>
<thead>
<tr>
<th>Table 6. Post-project Reflection Submissions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-reflection</strong></td>
</tr>
<tr>
<td>----------------------</td>
</tr>
<tr>
<td>Project 1</td>
</tr>
<tr>
<td>Project 2</td>
</tr>
</tbody>
</table>
I have always wanted to be able to express myself more clearly; to share my ideas with others more effectively; to make good arguments and illustrate my point; in short . . . to become a better writer. Until recently this was only wishful thinking, but now I feel a lot more confident about my writing! I can even see a positive change in my attitude towards taking up new writing projects. . . . Specifically, I learned to summarize, analyze, question, and evaluate written text. The following is an abstract from my Project 1 "Finding the "Sweet Spot”—Review of Daniel Coyle’s The Talent Code" (pg.1) Coyle's main argument is that by engaging in deep practice—the process of "struggling in certain targeted ways" (18) where you operate at the edge of your ability and are bound to make mistakes—you in fact become better. To support this argument the writer employs anecdotes and uses evidence collected through personal observation, as well as examples and studies conducted by Robert Bjork, chair of psychology at UCLA.

[Here is an example of where I have learned how to] summarize, question, analyze, and evaluate visual texts:
From page 1 [Project 2]:
The image is almost monochrome: black, white, and shades of cold grey-green. From the far left of the picture, the cold light of a computer screen streams outwards, spreading across the desk, and harshly outlining the profile of a bearded young man, casually dressed in a t-shirt. The blackness of the t-shirt merges into the enclosing darkness at the right-hand side of the picture.

In the foreground, on the desk, we can see the casually discarded lens-cap of the camera. From the angle of the picture, it seems that the camera itself has been placed on the corner of the desk and tilted upwards just enough to bring the face of the young man into view. This gives the impression of a spontaneous, improvised, picture, where perhaps a few moments were spent to sweep away some of the clutter. The untidy shelf of well thumbed documents in the background still shows that this is a real workspace.

At first glance this resembles those stock pictures of computer work, where a smartly dressed photo model sits calmly at a computer terminal, surrounded by the blue neon glow of imagined cyberspace. But in this portrait, the textures haven’t been photoshopped away, and the facial expression does not express cool robotic competence, but instead something much more familiar, much more human. This is the expression of concentration at a demanding task.
8, students indicated fairly strong confidence across most domains, especially in their abilities to draft, revise, edit, read critically, and summarize.17

Students indicated less confidence with reading and responding to other writers’ work, making them well suited for a MOOC because its reliance on peer feedback provides the occasion to hone these skills. Such data suggest that those integrating writing into MOOCs may do better to use peer-assessment feedback

<table>
<thead>
<tr>
<th>Table 8. Pre- and Post-course Self-efficacy Survey Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please indicate how certain you are that you can engage in each of the following:</td>
</tr>
<tr>
<td>Scale: 0 (Completely Certain Cannot Do It) - 10 (Completely Certain Can Do It)</td>
</tr>
<tr>
<td><strong>The Writing Process</strong></td>
</tr>
<tr>
<td>Draft a piece of writing</td>
</tr>
<tr>
<td>Workshop writing</td>
</tr>
<tr>
<td>Revise my writing</td>
</tr>
<tr>
<td>Edit my writing</td>
</tr>
<tr>
<td>Respond to the work of others in order to help them revise it</td>
</tr>
<tr>
<td><strong>Higher-order Learning Objectives</strong></td>
</tr>
<tr>
<td>Read a written text critically</td>
</tr>
<tr>
<td>Analyze and evaluate a written text</td>
</tr>
<tr>
<td>Read a visual text critically</td>
</tr>
<tr>
<td>Argue and support a position in writing</td>
</tr>
<tr>
<td>Recognize disciplinary expectations</td>
</tr>
<tr>
<td>Transfer writing skills to new writing contexts</td>
</tr>
<tr>
<td><strong>Smaller-order Learning Objectives</strong></td>
</tr>
<tr>
<td>Write concisely</td>
</tr>
<tr>
<td>Create an effective introduction</td>
</tr>
<tr>
<td>Cite the work of others in my writing</td>
</tr>
<tr>
<td>Write a strong conclusion</td>
</tr>
</tbody>
</table>

Note: Table contents are abridged. See Appendix K for full data.
than AES—which provides only a computer evaluation—in terms of meeting student interests. Students were also slightly less confident with transfer and adjusting writing to different learning situations. Here, too, we imagine that the multidisciplinarity and heterogeneity of the MOOC participants provided a strong context for students to advance these skills.

Responses also indicated slightly less confidence regarding smaller-order learning objectives, such as writing introductions and conclusions. Since most enrollees were professionals with degrees, they may have been looking for instruction in these aspects.

Post-course self-efficacy survey data indicated gains in confidence across all domains, even when the data were correlated for the same individuals. The largest reported learning gain was in responding to the work of others to help them revise. Similar gains were reported in transfer, adjusting writing to different contexts, and in writing conclusions and introductions. The most minimal gain was in workshopping writing, which is not surprising as Hangouts space was limited.

Post-course surveys were sent to every enrollee, regardless of engagement or completion. Tables 9 a-c show general satisfaction across several key domains, which is expected given that less satisfied learners presumably would have stopped participating and may have also been less inclined to complete a survey. 18

Although 1,000+ respondents is a comparatively larger number of people than typically found in a first-year writing class, this N is an exceedingly small percentage of the overall enrollees, approximately .012 percent. A challenge in working at this scale is deciding whether to interpret such data as representative or unique.

<table>
<thead>
<tr>
<th>Table 9a. Post-course Survey Selected Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>How much do you agree with the following statement? Scale: 1 = Strongly Disagree - 5 = Strongly Agree</td>
</tr>
<tr>
<td>For the amount of time I invested in this course, I’m happy with what I learned.</td>
</tr>
<tr>
<td>I found this course personally fulfilling</td>
</tr>
<tr>
<td>I learned what I was hoping to learn in this course</td>
</tr>
</tbody>
</table>
Table 9b. Post-course Survey: Overall Experience (N = 1009)

<table>
<thead>
<tr>
<th>Rate your overall experience with the course (1 = Poor; 7 = Excellent)</th>
<th>Mean 5.59</th>
<th>Standard Deviation 1.41</th>
</tr>
</thead>
</table>

Table 9c. Sample Post-course Survey Comments

This has been food for the hungry to me. I learned not only about writing, but about academic protocols for writing. When I signed up for the course I had no idea how nicely it would fit into the area of my learning that I want to develop. I want to be able to participate in discussion about the economics of this world, and this class has given me several invaluable tools to do this. THANK YOU, well done!

I undertook this course after encouraging 60 of my students (mostly 17 year olds) to sign up with me. I teach English to pre-university students at a government school. Most dropped out along the way but I think about 10 completed (I’ve yet to get all their feedback as we’re on break now). This was a brilliant addition to our class activities and an excellent introduction to university study for my students.

**Intensive Holistic Portfolio Rating**

Held in June 2013 and led by assessment expert White, this portfolio rating, conducted by an expert team of nine professional writing instructors and writing-center tutors trained as raters, sought to answer the following questions:

- How many of the students completing EC have achieved the learning objectives of the course and to what degree?
- What is the interrater reliability (IRR) between expert raters and MOOC peer raters on Project 3 (Case Study) and Project 4 (Op-Ed)?

Learner portfolios, which consisted of a sample set of 250, included the following:

- Reflective Cover Letter/Final Essay
- Project 1 Draft and Final Version
- Project 2 Draft and Final Version
- Project 3 Draft and Final Version
- Project 4 Final Version
The sample set was derived by correlating those who had completed the course with those who had agreed to participate in the IRB study; this correlation yielded 315 portfolios, from which we drew a random sample of 250. Staff in Duke University’s Center for Instructional Technology downloaded data from the EC site, de-identified it by discarding email addresses and names, and then added numerical identifiers of numbers 1–250. Of these, three portfolios contained damaged files, so our final sample size was 247. De-identified data were placed in folders on a specially created Sakai website, Duke University’s learning management system.

We designed assessment measures on established findings for assessment that identify holistic and primary trait analytics as closely related and corroborative (Breland and Gaynor; Spandel and Stiggins; Veal and Hudson; White, “Holisticism”). Scoring rubrics were developed in draft form; then a subset of our expert raters tested and revised them with a sample of portfolios. We held a half-day norming session with the entire team, where we slightly revised the rubrics, followed by two and a half days of rating. (See Appendix H for Portfolio Assessment Rubrics.)

We asked our assessment team to record three different scores for each portfolio:

1. A holistic score for the overall quality of the portfolio, based on a careful reading of the reflective cover letter, supplemented by skimming the cited passages in the portfolio to find the passages cited in the cover letter in order to verify or discredit the claims made in the letter (six-point scale) N=247;

2. A holistic score assessing progress toward the learning objectives for Project 3, using the same evaluative rubric peers used in the MOOC (six-point scale) N=95; and

3. A holistic score assessing progress toward the learning objectives for Project 4, using the same evaluative rubric peers used in the MOOC (six-point scale) N=95.

Each portfolio was read independently by two different raters. In cases where they did not have exact or adjacent scores, a third expert rater read for “adjudication” or “tertium quid.”

Some scholars challenge adjudication, suggesting that it unfairly increases IRR and eliminates important data points (Cherry and Meyer). For an expert
reading of MOOC student projects, we felt that expert adjudication was highly warranted in cases where raters had a discrepancy because we were primarily interested in arriving at as “expert” a reading as possible. However, because disagreement exists in assessment literature, we present results with and without tertium quid. Our intensive holistic portfolio rating demonstrated that MOOC writers can achieve learning objectives. Writing produced by learners was generally average to strong, though some selections rated lower. Table 10 shows that the mean score was over 4.20 Figures 5 and 6 show that approximately 64% of learners created portfolios that scored in the upper half of our scoring rubric.21 Of particular significance is that we learned from reading the portfolios that learners produced real responses to rigorous writing assignments. We were genuinely impressed with how seriously so many learners in our sample pool took their writing.

Given the confidence reported on the self-efficacy surveys, we might have expected a stronger overall performance on the portfolios. This suggests perhaps a disconnect between students’ perceived and actual writing abilities. Still, we were excited to see how much earnest writing was being produced through our portfolio reading, that MOOC writers can achieve learning objectives, and that writing learning gains are present for some MOOC writers.

One aspect we were especially interested in learning more about was the correlation between peer and expert assessors, and so we conducted comparisons across two projects. Literature on student peer assessment suggests a strong correlation between peer and expert assessment, especially when

---

**Table 10. Holistic Portfolio Scoring Mean and Inter-rater Reliability**

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Mean Score, without Tertium Quid</th>
<th>Inter-rater Reliability, without Tertium Quid</th>
<th>Mean Score, with Tertium Quid</th>
<th>Inter-rater Reliability, with Tertium Quid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Holistic Score</td>
<td>4.09</td>
<td>.683</td>
<td>4.2</td>
<td>.916</td>
</tr>
<tr>
<td>(1= Poor; 6=Excellent); N=247 (494 scores)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
peers make a “global judgment based on well understood criteria” (Falchikov and Goldfinch 315). Our data confirm this; Tables 11 and 12 show that expert and peer scores were relatively close, with peers being slightly more generous than experts. (See Appendix M for expert grade frequency and distribution across Projects 3 and 4.)
Table 11. Peer-Expert Rating Comparison and IRR, without Tertium Quid
(Six-point scale: 1=Poor - 6=Excellent; see Appendix H for rubrics)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Peer Mean Score (285 scores)</th>
<th>Peer IRR</th>
<th>Expert Mean Score (190 scores)</th>
<th>Expert IRR</th>
<th>Difference between Peer and Expert Average Scores</th>
<th>Peer-Expert IRR Based on Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project 3</td>
<td>95</td>
<td>4.66</td>
<td>.529&lt;sup&gt;c&lt;/sup&gt;</td>
<td>4.56</td>
<td>.736</td>
<td>Peers .1 higher</td>
<td>.749</td>
</tr>
<tr>
<td>Project 4</td>
<td>95</td>
<td>4.63</td>
<td>.426</td>
<td>4.15</td>
<td>.616</td>
<td>Peers .48 higher</td>
<td>.596</td>
</tr>
</tbody>
</table>

a. The N for expert ratings is 95, and the N for peer ratings is 97. Two of the files were corrupt in the expert reading.
b. We used the same method to calculate these ratings as well: ICC.
c. This is calculated using the same method: ICC. We used the first three scores used for a project. Some projects received more than three scores, but not all.

Of particular note are the IRR scores. In a quantitative meta-analysis of forty-eight peer-assessment studies, Nancy Falchikov and Judy Goldfinch found that “[t]he mean correlation over all the studies was 0.69, indicating definite evidence of agreement between peer and teacher marks on average” (314). While our IRR yield similar rates, the IRR for peer ratings is indeed lower than the expert IRR.
However, the IRR overall between peers and experts is about the same as the overall IRR between experts.

IRR data differ, though, across Projects 3 and 4. For Project 3, the IRR between peers and experts (.749, without *tertium quid*) suggests peer assessment for grading functions with relatively similar reliability as expert assessment. For Project 4, though, the IRR of .596 offers less confidence in peer-expert agreement. We hypothesize this may be due to the nature of Project 4, which was an op-ed, and students may have been unfamiliar with the genre. Research suggests that rating task has a high impact on the reliability of expert and lay readers (Schoonen, Vergeer, and Eiting); therefore, rating tasks in MOOCs should be considered when determining assessment.

Based on the comparatively small difference in mean grades assigned between peer and expert raters (with and without *tertium quid*), we learned not only that the rubric was reliable and that peers and experts have relatively close consensus, but also that peer assessment works on average as well as expert response for grading.

Such a conclusion, though, must be qualified: this means that peer and expert grading, under these circumstances and with this rubric, had high correlation. It does not necessarily mean that formative feedback has high correlation between experts and peers. That is, our peer-expert grading correlation does not indicate the degree to which peers can help one another learn and grow as writers. In fact, the unevenness in formative feedback suggests that MOOC writers cannot rely consistently at this point on MOOC peers for formative feedback. Perhaps this is one important difference between peer assessment in other contexts and in a MOOC. Still, we were pleased to see that the peer assessment in evaluative judgment proved to be overall nearly as reliable as expert assessment.

**Further Research**

MOOCs present many opportunities for future writing-related research. We see the following as especially critical:

*Peer Feedback (Formative).* What are the best practices in formative feedback in MOOCs, and how can we better support students in making use of feedback? Given the high potential from providing and receiving peer feedback, this seems a crucial area for more research.

*Peer Assessment (Evaluative).* More research is needed across genres, tasks, and disciplines on the reliability and validity of MOOC peer as-
assessment. Also, how should we share this research with MOOC enroll-ees so they can better understand the theory and practices informing peer assessment? Until this research is conducted and analyzed, any consideration of awarding college credit for completion of a writing-based MOOC is premature and unwarranted.

*ESL and Writing-based MOOCs in English.* Much discussion occurred throughout EC about interactions between ESL and non-ESL writers. More research is needed on the experiences of ESL MOOC learners and on interactions between learners across linguistic and cultural backgrounds, including world Englishes.

*Learner Analytics.* What motivates people to take a writing-based MOOC and how can their interests be served? How do demographics impact learner success? Such research could build on the excellent work already conducted in this area of online writing instruction (de Montes, Oran, and Willis; Hawisher and Selfe; Thatcher).

*Writing Pedagogy.* What is the impact on faculty of teaching writing in a MOOC? How can a MOOC create learning communities? Can MOOC writing faculty integrate sequencing and discrete writing units? Scholarship is emerging in this area (Comer, “Learning”; Kolowich, “Professors”; Krause and Lowe).

*WAC, WID, and MOOCs.* How can writing be integrated into discipline-based MOOCs? For a beginning to this conversation, see Comer’s “MOOCs Offer Opportunities for Writers.”

*Attrition.* Why do people stop taking a MOOC? Who is more likely to disengage? Research in this area might also examine alternative ways of measuring student involvement since many MOOC learners enroll for purposes other than course completion.

*Blended Possibilities.* What elements from MOOCs might be shared with and adapted for blended use in more traditional classroom settings? Under what circumstances?

*Academic Honesty and Plagiarism.* How do we protect against plagiarism in a MOOC? We found several instances of plagiarism. As in face-to-face settings, it is difficult to differentiate between cheating and lack of familiarity with US academic writing citation conventions. Some work has already been done in this area (Conrad; Meyer and Zhu).
Conclusion: Success or Failure?
When facing disruptive innovations, it is tempting to revert to extreme concluding statements such as: Almost all students in EC failed to achieve learning objectives; it is impossible to teach writing in a MOOC. Or, EC was wildly successful, and we can teach writing meaningfully in a MOOC. Our data could offer evidence for either claim. Given the MOOC scale, it is highly unlikely that we could ever definitively conclude one way or another.

The truth is, for many learners the MOOC did not work; these learners found the format impersonal, received unsatisfactory peer feedback, and did not advance as writers. For some learners, it did work. They found the space and time to reflect on themselves as writers, engage with others about writing, share their writing with a global readership, and make gains in writing skills. The MOOC was likely better for some learners than whatever other educational alternatives were available to them; for many others, the MOOC was likely an inferior educational choice.

A great many others found some aspects meaningful to their growth as writers, and others less so. Thousands of people thought about themselves as writers or interacted with others on the forums about writing. Many learners made connections with other writers; writing projects provided rigorous writing opportunities, and many writers took their work for the course seriously. Even if some enrollees worked on their writing for only a week, or a day, they may have done more with writing than they otherwise would have. A total of 1,289 people earned a Statement of Accomplishment. Comer teaches twelve first-year students a semester at Duke University in her academic writing seminar (as director of first-year writing her other time is spent with administrative, service, and research responsibilities); it would take approximately fifty-three years for her to reach 1,289 learners.

MOOCs are likely to persist in some form, whether disciplinarily situated, blended, professionally focused, or otherwise. We need to learn more about how, when, and where writing can be integrated into MOOCs, and about how to tailor large-scale writing pedagogy for a vast range of learners. MOOCs cannot and should not replace face-to-face instruction. However, they are likely to be one of a number of postsecondary educational initiatives that will (for good or bad) disrupt and reshape higher education in the coming decades. Responding
with anger or passivity are two choices, but we must also conduct research. Scholars invested in writing studies must lead efforts to shape the role of writing, writing pedagogy, and writing assessment in MOOCs. If we do not, others who are likely to be much less invested in the values and best practices our organizations have established will certainly do so.

Perhaps of most importance, though, is that the MOOC suggests we revisit assessment at its core: What are we measuring and why? Surely we want to continue assessing within established frameworks. And to those ends, evaluation of student portfolios seems the best way to measure the success of a MOOC in helping students improve their writing. Our holistic portfolio rating suggests that much of the student writing in this sample was successful. Formative peer feedback was uneven. Peer evaluative assessment produces results similar to expert evaluations, though writing assignments play a significant role in this correlation.

Our experience in MOOC writing assessment, however, also sponsored an even more important question: Who should hold the responsibility for and privilege of assessment design, implementation, and analysis? While we feel strongly that these are core faculty responsibilities, we must acknowledge that strong pressures for outside evaluation exist and less-qualified evaluators stand ready to fill the vacuum if faculty do not undertake careful and credible research authenticating faculty assumption of those responsibilities. We see our work contributing to ongoing efforts in our field to tell data stories with a mixed-methods approach, offering a quantitative and qualitative narrative for data-driven inquiry and emphasizing the social construction of information (Honig and Coburn 592). MOOCs pose a question that researchers in many fields now struggle with: How can we best understand “Big Data”?

The ability of MOOCs to generate a tremendous amount of data opens up considerable opportunities for educational research. edX and Coursera, which together claim almost four and a half million enrollees, have developed platforms that track students’ every click as they use instructional resources, complete assessments, and engage in social interactions. These data have the potential to help researchers identify, at a finer resolution than ever before, what contributes to students’ learning and what hampers their success. (Breslow et al.)

Our experience suggests that Big Data has limited potential without also considering qualitative data.

Perhaps, for all the allure of “scale,” clickstream data, and learner analytics, what the research and assessment communities should really be looking at
are individual students. Certainly, Big Data assessment remains valuable, but
our experience of the MOOC suggests that Big Data must be fused deliberately
with a more individualized, learner-driven, and learner autonomous approach
toward assessment.

Openness and scale in the form of a MOOC suggest that students are not
only one component of assessment, but the primary part; students hold the
authority to define, construct, and disrupt their own frameworks for assessment. Their assessment of their own work through self-reflection and peer feedback become the primary measures of value. In this way, even within a
course of 82,820 learners, every single individual’s experience matters in the
totality of assessment. As much as we might develop learning objectives and align course materials to these learning objectives, MOOC learners bring their
own priorities, objectives, and interests to the writing MOOC. If making connections mattered for some learners, then their assessment should ask if they accomplished that and to what degree; if their priorities were watching videos,
gaining a career credential, becoming more equipped to teach their children
literacy skills, participating in lifelong learning, or fostering creativity, then
these are the terms that their assessment should measure.

In a somewhat surprising twist, then, given the scale, writing assessment in
the MOOC context functions as one of the most limited of rhetorical situations,
what Lloyd Bitzer calls an “exigency.” You write something to some audience,
even if it is only yourself, as in a diary, for some purpose. Assessment in a MOOC,
and perhaps throughout education, only makes sense in terms of accomplishing
what you want to or agree to achieve in a specific rhetorical situation.

Thus, our experience creating a hybrid MOOC, one that provided cMOOC
pedagogy in an xMOOC platform, also created our hybrid approach to assess-
ment: one that operated at scale, emphasizing connectivity, but relied above all
on what individuals chose to do with course material. MOOCs, though in many
ways too instructor-centered, underscore the need to place assessment in the
hands of learners themselves, explicitly encouraging each learner to construct,
develop, and experiment with assessment. The physical presence of the teacher
is not always as important for student learning as providing clear assignments,
the opportunity for many responses, requirements for revision, and reflective
writing. Therefore, portfolio assessment, when based on metacognitive student
self-assessments and focused on this combination of personal and course goals,
stands out as the most valid and reliable measurement technology.

This effort—paradoxically creating autonomous space for the individual
within a massive scale—seems not only the optimal approach to MOOC assess-
ment, but also for non-MOOC settings. In writing this, and in fact in teaching a MOOC at all, we see our efforts as part of a larger trajectory, so deeply embedded in writing studies, to participate in conversations about how we can transfer and adapt what we learn from one context about writing pedagogy, assessment, student writing, and higher education to others. We learned that well-planned assignments are a key element for a writing course or for any measure of the success of such a course. But, beyond that, assessment lies in the hands of learners. Peer feedback, with clear criteria, early response, and as much response as possible from multiple readers is almost as useful to students as expert response and is crucial for student learning.

One answer, then, to the question many are asking, “How can research into MOOCs contribute to an understanding of on-campus learning?” (Breslow et al.), may, after all, be one that urges faculty across disciplines to scale assessment itself to be more open and accessible, not only to the masses, but to individual learners themselves.

Acknowledgments

We would like to thank the members of our June 2013 assessment team: Calina Ciobanu, Benjamin Gatling, Heidi Giusto, Gale Greenlee, Patrick Horn, Lauren Spohrer, Jay Summach, and Katya Wesolowski. Jennie Saia provided event planning and implementation. Thank you to Yvonne Belanger, Matt Serra, Jessica Thornton, and Quentin Ruiz-Esparza for their statistical analyses. The EC team at Duke provided much expertise, feedback, and support: Kendra Atkin, Elise Mueller, Shawn Miller, and Lynne O’Brien. Appreciation goes to Marcia Rego for developing the self-efficacy surveys, and to the three disciplinary experts affiliated with English Composition I: Maral Erol, David Font-Navarrete, and Rebecca Vidra. Thank you also to the members of our unofficial consortium of writing course Gates Foundation MOOC grantees for their ongoing collaboration: Rebecca Burnett, Kaitlin Clinnin, Susan Delagrange, Scott DeWitt, Andy Frazee, Kay Halasek, Karen Head, Pat James, Ben McCorkle, Jen Michaels, Cynthia Selfe, and Louis Ulman. Funding for this assessment was provided by the Bill & Melinda Gates Foundation, the Duke University Thompson Writing Program, and the Duke University Center for Instructional Technology.

Notes

1. Anne Ruggles Gere et al. outline of the history of “White’s law”: “Originally posted on the Writing Program Administrators listserv, December 7, 1996, the entire quote reads: ‘I give you White’s law the truth of which I have noted for over..."
twenty years: Assess thyself or assessment will be done unto thee” (630). White attributes the nomenclature of “assessodynamics” to the following: “Keith Rhodes (in conversation) has named my little proverb on this matter ‘White’s first law of assessodynamics’ (“Misuse” 33).

2. There is much literature describing MOOCs and their history. Good starting points might be the Chronical of Higher Education’s “What You Need to Know about MOOCs” or Educause’s “What Campus Leaders Need to Know about MOOCs.” Annotated bibliographies are also a good source: see Tina Christner’s bibliography and “MOOCs” from the University of British Columbia website.

3. For more on the Bill & Melinda Gates Foundation’s efforts regarding education, see “Postsecondary Success.” For critiques and implications of their approach, see Kovacs and Christie; Picciano and Spring; and Rhoades, Berden, and Toven-Lindsey.

4. EC also included significant support from Duke University and Coursera. Three other writing-based MOOCs were also awarded separate funding through the Gates Foundation: Writing II: Rhetorical Composing (Delagrange et al.); First-Year Composition 2.0 (Head); Crafting an Effective Writer: Tools of the Trade (Barkley, Blake, and Ross).

5. At the time we submitted the final version of this article, EC had concluded its fourth iteration, running June 26–August 27, 2015, with a total of 69,622 enrollees. The third iteration of EC ran September 22–December 22, 2014, with a final count of 58,369 enrollees, and the second iteration of EC ran April 21–July 14, 2014, with a final count of 92,660 enrollees.

6. According to a pre-course survey (N=9584). Ages: Under 18 (2.02%), 18–25 (27.15%), 26–34 (34.56%), 35–44 (18.75%), 45–54 (10.76%), 55–64 (5.21%), 65 or over (1.38%). Highest degree achieved: Less than high school (2.06%), secondary (8.89%), some college (14.67%), bachelor’s (36.95%), master’s (29.37%), and doctorate (4.06%). Employment: Precollege student (3.92%), undergraduate (15.09%), graduate student (17.17%), research scientist (6.24%), professional (26.19%), academic/professor (6.89%), teacher (10.29%), working full-time (37.31%), working part-time (13.34%), other (11.15%). The full-time percentage was calculated by eliminating all responses indicating student and other.

7. The three disciplinary consultants, Rebecca Vidra (natural sciences), Maral Erol (social sciences), and David Font-Navarrete (humanities) provided “spotlight” videos on writing in the disciplines (i.e., using images, public scholarship), participated in the forums, and facilitated several Google Hangouts workshops.

8. The course description, learning objectives, assignment details, and other course details are reproduced, drawn from, or abridged from the EC website (Comer, EC).

9. For more on the distinction between xMOOCs and cMOOCs, see Yuan and Powell.
10. See Appendix C for a table comparing these objectives with the “WPA Outcomes Statement” and the Framework for Success; see Appendix D for unit learning objectives.

11. There have been some public misperceptions about the required page lengths of these major projects. For instance, in “The Rise of the Online Writing Classroom,” June Griffin and Deborah Minter cite the total required pages for the entire course as ten pages, in part to make an argument that the course did not have rigorous writing requirements for students (150). Griffin and Minter may have tabulated these page lengths from the initial EC advertising page. However, this page did not specify whether page lengths were single- or double-spaced since at the time we were still in conversation with Coursera about acceptable page lengths for their platform. Actual word counts within the course are indicated here in order to show that students completing EC wrote at least 2,700 words, some 3,600, for their major projects, which is more like fifteen pages double-spaced. These major writing projects also went through an extensive drafting and revision process. Students also wrote considerably more within the course in terms of page equivalence if we take into account post-project self-reflections, the final reflective essay, peer feedback, and forum contributions.

12. Additional assessments also included two instructor focus groups (one public and recorded, one private) to discuss how teaching the MOOC impacted writing pedagogy.

13. Those who submitted a final version may or may not be the same individuals who submitted a draft. The final version assignment portal was open to anyone enrolled in the course. We did not at this time run a cross-correlation to determine by user ID how many of those submitting a final version were the same individuals who submitted a draft.

14. For more examples of peer feedback, see Appendix J.

15. Drafts were scored 0 for incomplete and 1 for complete. Average scores earned on drafts were as follows: Project 1: .91; Project 2: .91; Project 3: .91. The average score reflects the average score of the peer reviewers after the lowest of the four peer evaluation grades were eliminated and the remaining ones averaged together within the Coursera platform. On Project 1, the raw average score (the average prior to eliminating the lowest of the four scores) received on our scale of 1–6 was 4.47; however, the refined average score (the average after eliminating the lowest of the four scores) received was 4.51. Project 2: raw=4.51; Project 3: raw=4.55; Project 4=4.43. Sometimes learners received slightly fewer or more peer evaluations depending on how many learners participated in the peer evaluation process.

16. Although one would assume that most people completing a post-project reflection had actually submitted a final version of a project, these may or may not be
the actual people who submitted final versions of the projects. We did not at this time run a cross-correlation by user ID to see how many of those who submitted a post-project reflection were also those who submitted a final version of a project. The post-project reflections were open to anyone enrolled in the course.

17. These pre- and post-course self-efficacy surveys were developed by Marcia Rego, director of assessment in the Duke University Thompson Writing Program. The N figures in the table are averages of the N for each specific question. N fluctuates slightly for each question since not all respondents answered every single question. Pre-course N varies 10258–10436; post-course N varies 919–932; correlated N varies 470–484.

18. Of those who responded, 67% indicated that they had earned a Statement of Accomplishment and 33% indicated that they had not earned one.

19. This process is one of several options for resolving score differences. Robert Johnson, James Penny, and Belita Gordon, who have conducted a review of score resolution methodology, offer a clear explanation of the process we employed:

   When discrepancies in ratings occurred, a majority of education agencies required an adjudication process in which an expert rater, who was also referred to as a table leader or adjudicator, independently assigned a third score. After obtaining this third rating, agencies applied a score resolution rule to form an operational score from the three reviewers' ratings. Although a myriad of such score resolution methods were identified, resolution most often involved (a) combining the score of the expert with the score of the closest rater or (b) replacing the scores of the raters with the score of the expert. (231)

For more information about adjudication, also known as “arbitration or moderation” (Johnson et al., “Score Resolution: Investigation,” 301), see Johnson, Penny, and Gordon’s “Score Resolution and the Interrater Reliability of Holistic Scores in Rating Essays.”

20. We calculated the interrater reliability (IRR) using Intraclass Correlation (ICC), in which the subscript 1 references a one-way random assignment and subscript k references the mean of the ratings. There are many methods for determining IRR, but our choice reflects the consultation of assessment experts at Duke University: Jessica Thornton, Matt Serra, and Quentin Ruiz-Esparza. For more on ICC, please see Richard Landers.

21. Given that 99% of our portfolio learners earned a Statement of Accomplishment in the course (see Appendix J for grade details for our portfolio sample pool), this figure of 64% might seem odd. Why did 36% of students earn a Statement of Accomplishment when their writing was unsuccessful? However, students’ course grades were based on elements outside of the final writing projects: 55% of the
course grade was derived from elements such as draft completion, peer feedback, and self-reflection; 45% of the grade came from the final versions of the four major writing projects (see Appendix B for syllabus). This kind of grade calculation occurs in many writing classes, which can include course credit for such aspects as participation, attendance, homework, drafts, verbal presentations, peer feedback, and quiz responses. Perhaps, though, since in a MOOC these other aspects have less direct oversight, the final course grade should be anchored more firmly in the final versions of writing projects.

22. Comer, Charlotte Clark, and Dorian Canelas have recently completed research that addresses this in part through EC and an Introduction to Chemistry MOOC (Comer, Clark, and Canelas). Comer and Canelas were successful recipients of funding from the MOOC Research Initiative, a competitive grant competition run by Athabasca University and George Siemens, and funded by the Bill & Melinda Gates Foundation.

Appendixes


Works Cited


de Montes, L. E. Sujo, Sally M. Oran, and Elizabeth M. Willis. “Power, Language, and Identity: Voices from an Online Course.” Computers and Composition


Herrington, Anne, and Charles Moran. “What Happens When Machines Read


Mabrito, Mark. “Electronic Mail as a


Selfe, Cynthia. “Technology and Literacy: A Story about the Perils of Not Paying Attention.” College Composition and


**Denise K. Comer**

Denise K. Comer (sites.duke.edu/denisecomer/) is an associate professor of the practice of writing studies and director of first-year writing at Duke University. She teaches face-to-face and online writing courses, and a MOOC. She earned the 2014 Duke University Teaching with Technology Award. Her scholarship has appeared in leading composition journals and explores writing pedagogy and writing program administration. She has written a textbook on writing transfer, *Writing in Transit* (2015), and a dissertation guide, *It’s Just a Dissertation*, co-written with Barbara Gina Garrett (2014). She has also designed and developed a Web-based course, *Writing for Success in College and Beyond* (2015).

**Edward M. White**

Edward M. White is a visiting scholar in English at the University of Arizona and professor emeritus of English at California State University, San Bernardino, where he served as English department chair and coordinator of the UWP. He was coordinator of the statewide CSU Writing Skills Improvement Program, directed the WPA consultant-evaluator service, and served two terms on the CCCC executive committee. He has authored over 100 articles and book chapters on literature and the teaching of writing, as well as five composition textbooks, and has coedited four essay collections. A second edition of his *Teaching and Assessing Writing* (1994) received an MLA Mina Shaughnessy award “for outstanding research.” His work was recently recognized by *Writing Assessment in the 21st Century: Essays in Honor of Edward M. White* (2012), and by the 2011 CCCC Exemplar Award. His latest book (with Norbert Elliot and Irvin Peckham) is *Very Like a Whale: The Assessment of Writing Programs* (2015).