Collaborative Ecologies of Emergent Assessment: Challenges and Benefits Linked to a Writing-Based Institutional Partnership

This essay reports on a writing-based formative assessment of a university-wide initiative to enhance students’ global learning. Our mixed (and unanticipated) results show the need for enhanced expertise in writing assessment as well as for sustained partnerships among diverse institutional stakeholders so that public programming—from events linked to classroom-level learning to broader cross-unit mandates like accreditation—can yield more rigorous, responsive, and mixed-method assessments.

This essay reports on a writing-based formative assessment of a university-wide initiative to enhance students’ global learning. Our mixed (and unanticipated) results show the need for enhanced expertise in writing assessment as well as for sustained partnerships among diverse institutional stakeholders so that public programming—from events linked to classroom-level learning to broader cross-unit mandates like accreditation—can yield more rigorous, responsive, and mixed-method assessments. Below, we describe the university-wide Quality Enhancement Plan (QEP)

CCC 69:2 / DECEMBER 2017

Copyright © 2017 by the National Council of Teachers of English. All rights reserved.
that framed our own multifaceted partnership; characterize one specific classroom's course context of student writing and that instructor's role in our project; and describe the informal network of classrooms that collaborated in hosting a scholar-writer's visit and evaluating its results. We also revisit the two-step process employed by two graduate student researchers for using student writing to assess the particular initiative's impact; and we suggest implications for future research on writing assessment situated within institution-level programming. We hope our work is helpful for teacher-scholars seeking to improve their pedagogy through classroom-level assessment exercises, but also for those wanting to get involved in institutional partnerships for collaborative assessments.

Our project report does not describe a formal study of writing assessment with a front-loaded research design. Rather, this essay revisits a partnership that brought together several leaders of writing-focused humanities programming with a university’s associate provost for institutional effectiveness to consider together whether and how student writing might provide a pathway to evaluating activities linked to the QEP developed for regional accreditation. At a micro level, faculty members involved tried out classroom-level writing assessment to enhance their future pedagogy and their ability to report on teaching in ways that are becoming essential for acquiring curricular resources (in this case, funds to support a visiting scholar-author). On a larger scale, by carrying out this assessment project early in the multiyear cycle of the QEP, this essay’s coauthors hoped to provide adaptable models for other instructors and to help the university-wide QEP committee identify ways for tracking progress in its global learning agenda.

As a meta-assessment exercise, the work presented here shares several findings. Methodologically, we demonstrate the value of combining quantitative and qualitative approaches to create portraits of learning achieved through humanities-oriented programming. In that vein, our essay itself includes preliminary analysis of some data quantitatively, but it also employs narrative description and interpretation to situate student writing from the project within a particular setting shaped by timing concerns, institutional and classroom cultures, and the values guiding teaching and research among the coauthors. Thus, we highlight the “collaborative ecologies” in which our assessment was embedded by blending quantitative and qualitative approaches while situating both forms of interpretation within
their local contexts. Kathleen J. Ryan, in “Thinking Ecologically,” emphasizes links between place-oriented epistemologies and “a compatible rhetorical ecological feminist agency” (75); Stuart MacMillan extols “[e]cological inquiry” that explores “learning and activity as interaction between individuals and environments” (346); and Kristie S. Fleckenstein et al. call for an “ecological orientation” that “conceives of activities, actors, situations, and phenomena as interdependent, diverse, and fused” together (388). For this report on our project, accordingly, we note specific ecological features of localized teaching and learning from the classroom level to the larger university-level context, while also taking into account the individual situations and interpretive approaches of the coauthors.

Pragmatically, our essay also illustrates the role writing can play in humanities-based programming and in related accreditation contexts. Current research on writing assessment analyzes the place of rubrics in assessing writing (Broad, What), the interplay between race and class in writing assessment (Inoue), and, recently, the theory and practice of large-scale assessments of writing programs (White et al.). Yet the role of writing in university-level program evaluations not directly tied to writing pedagogy has received less attention. This gap in the scholarship is problematic. First, many program evaluations, while not explicitly affiliated with writing instruction, still use student writing as the primary assessment tool. For instance, assessment experts advocate using “student work artifacts already embedded in courses” (such as writing assignments) as a way for various campus programs to “develop outcomes that reflect what is actually happening without adding much testing time” (Wehlburg 11). Second, considering, as Bob Broad reminds us, that assessments are “public records of rhetorical values” (What 9), investigating writing in multiple sites of assessment can illuminate how programs not exclusively tied to writing pedagogy construct rhetorical values about writing. And third, understanding the role of writing in assessment projects not explicitly tied to the discipline of writing studies can help us locate potential for interdisciplinary relationship building among faculty, administrators, graduate students, and undergraduates engaged in assessment.

For example, our assessment project was not tied to a single writing course, nor was it tied to a writing program, per se. The process we developed here was essentially tied to a specific humanities initiative involving a renowned author’s campus visit: we were assessing the efficacy of that
visiting scholar’s engagement with students as part of a recently launched university-wide QEP promoting students’ global awareness. From faculty leadership for the individual events to student and staff engagement to local university administrators tasked with managing the overall QEP agenda—assessing this initiative in any substantive way had to be multipronged. But perhaps the most salient part of the assessment for readers of this journal is that despite the project not being explicitly connected to a writing program or even writing pedagogy, writing was still the object and the means of assessment.

Through open-ended questionnaires administered in seven course sections that engaged with a visiting scholar, this essay’s authorial team analyzed written responses to produce a formative evaluation of this particular initiative within the evolving context of the QEP. Thus, one key takeaway from our work is that collaborative assessment projects like ours are important for discovering, promoting, and implementing significant intellectual exchanges across disciplines and programs. For writing studies scholars in particular, this collaboration also shows how formative writing assessment partnerships can have a meaningful impact on university-wide practices.

While building such collaborations may help institutions better assess their cross-disciplinary enterprises, we should still ask if working across such divides will be productive for composition/rhetoric teacher-scholars, as well as their students. Some compositionists have famously chastised those in psychometrics and educational measurement as treating students like numbers: “We don’t feel comfortable [talking to them],” Peter Elbow and Pat Belanoff wrote in 1997, “so long as they continue to worship numbers as the bottom line. We think teaching is more important and more interesting than assessment” (21). However, Brian Huot suggests that writing researchers need to do a better job learning the language of educational measurement, particularly terms like validity (how well a construct measures what it says it measures) and reliability (how consistent scores are across replications of an assessment). Huot argues that the better claims writing researchers can make about validity in writing assessment, the better writing researchers can engage with those in other disciplines or in educational measurement. “If we can promote the regular use of validity arguments that attempt to be compelling for all of those who work in writ-
ing assessment,” Huot writes, “then it might be possible to ease the current climate of isolation, since both camps in writing assessment would need to know about each other in order to make convincing arguments for validity” (56). We agree that seeking common language is important. Yet, we argue that flexible partnerships do not require full adoption of such goals as formal statistical significance as reported via t-tests, standard deviation, and other quantitative techniques. (See our reporting of data below, where we do give range.)

Connecting across these boundaries can bring benefits at multiple levels of exchange where writing program administrators (WPAs), composition instructors, and their students interact with others within university culture. On their side, upper-level administrators need to develop a clearer understanding of how teachers and program leaders in writing studies evaluate student learning—and why. On another side, through collaborations like ours, writing studies–focused team members can take away learning for designing and assessing interdisciplinary programming, as well as how to assess affective elements of learning in their own classrooms.

We should also note that assessment, as Michael F. Middaugh writes, “has become the primary tool for understanding and improving the ways in which students learn and for developing and enhancing those institutional structures and programs that support student learning” (x). Accreditation agencies, as Middaugh further observes, play a key role, and the “culture of evidence” that is part of such processes certainly brings both benefits and challenges to collaborative inquiry like ours. Thus, as Middaugh advises, we sought to “embrace the opportunity to measure student learning and institutional effectiveness as a vehicle for more effectively communicating,” to “both internal and external constituencies,” how an endeavor like a QEP can align with institutional mission (x–xi). Though fully recognizing the need to be able to tailor assessments to local needs, and even classroom-level learning objectives (Barlow et al.), our analysis below includes consideration of how pedagogical dimensions of our project were both supported and constrained by being situated within a university-wide accreditation agenda, through a rubric originally designed by the Association of American
Colleges and Universities (AAC&U). That is, the nature of the national push toward accountability and, frankly, the funds required to bring in visiting scholars necessitated the design of an assessment project aligned with broader national, regional, and institution-level frameworks.

**University-Level Context: The QEP at Texas Christian University**

Our regional accreditor, the Southern Association of Colleges and Schools Commission on Colleges (SACS-COC), requires a QEP from each institution as part of ten-year reaccreditation. When this project began, Texas Christian University (TCU) had chosen “Discovering Global Citizenship” as its thematic focus. TCU’s website describes this QEP as including “six initiatives designed to engage the TCU community with the world while providing international and comparative experiences for students” (“Who”). The institution’s overview clarifies these components of the program:

- Global Innovators (enabling scholar-activists to build long-term partnerships with TCU),
- Virtual Voyage (using technology to bring experts to the university),
- Local Global Leaders (connecting TCU with global communities in the region’s own population),
- Global Academy (involving students in collaborations with social justice projects in other nations),
- TCU Abroad (enhancements to Study Abroad), and
- Visiting Scholars (short-term visits to campus by global learning experts).

At the time our Visiting Scholars project was launched, formative assessment of the Discovering Global Citizenship program was gradually becoming multifaceted. Since the QEP was in its early stages, one approach for tracking impact was to straightforwardly tally the number of students attending an event or participating in any project. For instance, students would swipe university IDs into a card reader when entering a public talk. Longer-term plans were developing for more inductive measures, such as mining portfolios, over time, as part of a separate campus-wide agenda that invites students to create digital records of their co-curricular
and curricular learning starting in their first year at TCU, culminating in a FrogFolio (named to align with the university frog mascot) in their fourth year. With the QEP in its own first year as our project occurred, and with the portfolio initiative being piloted only with small groups of students, there was not a clear path to integrating our assessment effort with FrogFolio. However, FrogFolio’s focus on gathering artifacts and encouraging reflection did highlight possibilities for connecting student writing from our Visiting Scholar project with the university’s emerging effort to track student learning through writing. These features of the local context, along with language in the QEP’s mini-grant proposal directions, encouraged this essay’s coauthors to seek out the associate provost for institutional effectiveness as a partner in imagining writing-based formative assessment for our Visiting Scholar activities, spread over several days.

Fortunately for faculty seeking QEP mini-grants, Associate Provost Catherine M. Wehlburg regularly makes herself available to support assessment for individual projects. Under her leadership, TCU’s Office for Institutional Effectiveness aims to view assessment as “not only a source of evidence that demonstrates our effectiveness in achieving our mission,” but also, and “more importantly, a process and a tool for improving student learning” (“Institutional Effectiveness”). Specifically, in helping guide assessment of the QEP, Wehlburg regularly meets with faculty and staff who develop programming on global citizenship to facilitate data gathering for SACS-COC. In this case, the richness of our preliminary informal conversations about our Visiting Scholar project led to a more sustained partnership exploring student writing as a window into QEP-related learning.

**Classroom-Level Context: The Grant Proposal’s Use of Writing**

Our Visiting Scholar project’s principal investigator (PI), Sarah Ruffing Robbins, sought to bring Kenyan literary figure Ngũgĩ Wa Thong’o to campus in spring 2014, the first semester of the QEP. Ngũgĩ’s multiday visit connected to several of Robbins’s ongoing pedagogical aims. Curricular goals within one college on campus also shaped this project. In the Honors College, a “Cultural Visions” requirement calls for students to take two courses focused on global, cross-cultural understanding. Robbins had recently proposed a new “Cultural Visions” course informed by Mary Louise Pratt’s scholarship on “contact zones” and had created a syllabus she would be using for the first offering in spring 2014.
Robbins envisioned the new Contact Zones class as connecting curriculum with public programming to address social justice issues. In that regard, the particular Visiting Scholar initiative also tapped into her research interest in using writing to assess learning in “soft” qualitative fields of study and interdisciplinary public projects. Thus, in planning for her proposal, Robbins built on her prior experience in global learning, her research interest in documenting writing-to-learn’s pedagogical power, and her belief in the value of writing in cross-disciplinary work (Pratt, Imperial Eyes).

The Contact Zones class brought in several visiting experts. For example, a Skype “visit” by a Peace Corps worker described his time in Africa. Also, TCU professor Mona Narain presented research in postcolonial studies. In each case, the avenue for assessing student learning through such engagements was writing, including texts for group presentations, reflective self-assessments, and an online discussion board. At the classroom level, the course used writing to promote and assess intercultural learning. But Robbins was also seeking to capitalize on the new resources being made available through the QEP. Thus, she aimed to bring a “big name” global scholar to her class. The choice of Ngũgĩ was partly based on the applicability of his own writings to the QEP.

Overall, Robbins’s proposal for funding emerged from her sense that the QEP offered an opportunity for blending writing-oriented teaching/learning goals with a formative Scholarship of Teaching and Learning (SOTL) research opportunity. But the mini-grant proposal also needed to show how the visiting scholar would contribute to a university-wide program. So one challenge became how to integrate assessment goals for the class with the larger-scale assessment needs for the QEP. On the benefit side, thinking beyond the classroom level led Robbins to consider ways of gathering comparative student writing data from several different courses. So Robbins folded into the grant proposal a plan for having Ngũgĩ’s visit link up with courses in the English Department and the Honors College, both of which provided funding for the project.

One way the proposal structure shaped the assessment involved what the format (and the QEP language) presented as a “taxonomy” for any work to be supported with Discovering Global Citizenship funding. The proposal form noted:
The Global Citizenship Taxonomy builds from the first outcome up through increasingly more complex cognitive abilities and actions. That is, the taxonomy begins at a lower cognitive level approach to global issues and moves up through the development of an active participant in the global community with the abilities and desire to make responsible decisions about global issues.

Within that framework, here is the taxonomy of “Student Learning Outcomes”:

**Outcome 1:** Students will *identify* global issues from perspectives of multiple disciplines and cultures.

**Outcome 2:** Students will *discuss* critical questions about the impact of global issues on domestic and global communities.

**Outcome 3:** Students will *develop* cultural empathy and intercultural competence.

**Outcome 4:** Students will *make* responsible decisions about global issues.

As the taxonomy indicates, outcomes 1 and 2 address relatively low-level learning outcomes keyed to verbs that affirm that hierarchy (i.e., “identify” and “discuss”). In contrast, both “develop” and “make” (as well as their objects—“cultural empathy and intercultural competence” and “responsible decisions”) point to more productive and sustained higher-level critical thinking. Further, these outcomes are more consistent with the types of learning generally sought in the humanities, where the banking model of delivering content is neither the main goal nor something that would tend to be assessed through open-ended writing processes and products (versus, say, a multiple choice test of whether students could “identify” content). Yet, one factor that actually dissuaded Robbins from selecting outcome 3 in her proposal was the framework’s linking of “empathy” and “intercultural competence.” Although recent research on reading fictional narrative and on text-based engagements has underscored a related capacity to build empathy (see Berns et al.) and therefore “intercultural competence,” Robbins’s understanding of the taxonomy conjured up such highly complex learning as that measured by the Intercultural Effectiveness Scale (IES)—and thus an outcome that would be difficult to achieve within the time frame of the Nguĩ˘gi visit (“Intercultural”).
For individual students, the total visit experience would entail some mix of reading material by Ngũgĩ before his arrival, attending the public lecture, engaging with him during a class visit, discussing his work with others, and writing in response to these activities. But it would not likely include more sustained opportunities to apply that learning toward the kind of “critical empathy” called for through more intentional, longer-term activities such as those outlined by William DeGenaro in “The Affective Dimensions of Service Learning.” Here was a clear downside of involving multiple courses taught by multiple instructors who typically would be giving, at most, one or two days’ worth of class time to this project. Where this configuration strengthened the original grant proposal by promising substantial numbers of undergraduates to be impacted through the individual class visits and the public talk, it also diluted the possibility of generating student writing reflecting deep, sustained engagement with global learning through Ngũgĩ’s work. Ironically, however, we came to recognize that the “develop cultural empathy and intercultural competence” outcome would have been more in tune with this project’s activities, including its grounding in discussion of imaginative texts, its emphasis on collaborative conversation, and its use of writing to promote cross-cultural learning.

There is also a complicated interplay here among individual classroom ecologies, as well as between those classrooms and the larger QEP framework. As we show below, we issued the same prompt across seven course sections, asking students to reflect on their experiences with Ngũgĩ. However, there were different pedagogical considerations in each section: the texts students had read often differed; the level of student engagement with Ngũgĩ varied; even the way each instructor framed Ngũgĩ’s visit differed. In short, each instructor implemented his or her own pedagogical approach to the materials that preceded and immediately followed Ngũgĩ’s visit. If each instructor had employed the exact same lesson plan and assigned the exact same books, and if all students had then attended one “standard” in-person exchange with Ngũgĩ, we might be in a better position to generalize about the impact of his visit on student learning. But then again, even if that were the case, each student brings a myriad of diverse prior experiences and expectations, not to mention the fact that we would never be able to say for certain who actually read the books, or whether students actually grasped the contextual material of the talk in the same way. In other words, critical pedagogies are determined by situational, contingent ecological
relations and are effective only when those differences are embraced and nurtured instead of standardized for the appearance of replicability. Thus, one goal in this formative assessment enterprise was to make visible those differences and ultimately to ask how we might incorporate varied, nuanced, classroom-based learning and teaching differences into program assessment. The context of this project, then, includes honoring individualized instructor contributions as well as the institutional context of the QEP’s evolving approaches to program assessment.

Data Collection and Analysis

When TCU’s QEP was first designed, a faculty/staff committee prepared four rubrics corresponding to the four stated outcomes for student learning (identified above). Versions of these rubrics, originally created by the Association of American Colleges and Universities (AAC&U) to contribute to assessment practices in higher education, have been used across many institutions (Sorey and DeMarte; Pike). Having linked Ngũgĩ’s visit with the “Discussion of Critical Questions” outcome (outcome 2), this essay’s coauthors generated writing prompts that could be used across all courses hosting our visiting scholar. Our prompts emerged through dialogue among all four coauthors in regular meetings and email. While the lead faculty member and graduate students brought knowledge about the local teaching practices in humanities classrooms, the associate provost provided deep understanding of university-level assessment priorities and tools from professional organizations beyond TCU.

Consistent with that extra-institutional context, we applied the QEP learning outcomes as framed by the AAC&U rubric to construct the following questions:

1. Briefly describe the activities in which you participated in connection with the Ngũgĩ’ visit, either before, during, or after his time on campus. More specifically, what discussions did you participate in, and how? How did your involvement with the visit and its related activities impact or change the nature of your understanding
of his writing and/or the issues or themes he depicts there? (For example, what are some new perspectives that you are now considering?)

2. Reflect on the activities that you participated in before, during, and after Ngũgĩ’ Wa Thiong’o’s visit. Ngũgĩ’ discussed his writing and his lived experiences in Africa as they are tied to various global communities. What learning are you taking away from your participation in his visit? What more are you now hoping to explore in the future and how?

3. OPTIONAL: Teachers, feel free to add in your own instructor-framed question or questions based on your own selected readings, classroom activities, and/or targeted learning outcomes.

Faculty who agreed to participate in the Visiting Scholars program received copies of these prompts to distribute soon after Ngũgĩ’s time on campus. Most faculty gave out the prompts and collected responses during class, providing limited time (generally less than twenty minutes) for the task to be completed, with three of those six classrooms lacking on-site computers for student use. One faculty member had students complete responses outside of class but emphasized the exercise should not be given much time. Although all students were told their writing would be evaluated to assess the impact of the visit, they also knew that their texts would not be graded. Across participating courses, there were varying levels of commitment to the visiting scholar himself, with some instructors already familiar with Ngũgĩ and others being introduced to his writing for the first time; similarly, some faculty were much more committed than others to the QEP. These variations in standpoint likely impacted how different faculty members framed the exigency of the writing task.

Because students in different classes had been assigned different readings from Ngũgĩ’s oeuvre, the questions did not focus on specific content from any one text. There was significant variation in the pre-visit preparation, with some classes having read only a short story, essay, or nonfiction excerpt (e.g., “Minutes of Glory” or parts of Decolonizing the Mind) and others much longer narratives, whether a memoir (e.g., In the House of the Interpreter) or a novel (The Wizard of the Crow or The River Between). Students also had varying levels of direct engagement with Ngũgĩ. Some were
in classes Ngũgĩ visited in person, with those connections including both an informal presentation and Q&A. Others read one or more publications but did not experience a class visit. Furthermore, although all students on campus, including those not enrolled in the target classes, could attend Ngũgĩ’s lecture, none were required to do so. According to the student ID card-swipe, of the 155 undergraduates who participated in the classroom-level surveys, 125 attended the public lecture. There were also a number of other TCU and community members in attendance, including some from the local Kenyan immigrant community, which boosted attendance beyond the student tally. Given the varying connections with Ngũgĩ, another worthwhile analysis (beyond the scope of the current essay) would be to explore the degree to which differences in the students’ written responses correlate with their varying levels and types of engagement with the visiting scholar. Another avenue of inquiry could explore correlations between scores earned and students’ year in school: Robbins’s class, for instance, targeted first-year students, whereas one honors symposium and a technical writing course enrolled mainly third- and fourth-year students. Furthermore, from an institutional perspective, however useful it might be to have a record and interpretation of some students’ short-term responses to this particular visiting scholar, our analysis of this single set of writing samples cannot be counted on to have captured all the impact of this particular program on students’ longer-term global learning. Who, for instance, may have been inspired by this occasion to read more of Ngũgĩ’s writings, to select a Kenyan setting for travel abroad, or to enroll in a course, later, on African culture and history?

Viewing both our data and our methods for interpreting them as formative, we positioned our work as exploratory from the outset. Initially, our team of coauthors had been meeting weekly to prepare our rubric and hold broader conversations about writing assessment. After collecting student responses from seven instructors and establishing basic norming procedures, two graduate students, Tyler S. Branson and James Chase Sanchez (also coauthors of this essay), analyzed student writing from each section that interacted with Ngũgĩ. As mentioned above, in total we reviewed responses to the writing prompts from 155 students across seven sections. Three sections produced handwritten responses, and four used computers. In general, the handwritten responses were shorter than the typed responses. Many handwritten responses did not extend past fifty
words, and some typed responses were as long as two hundred. Consistent with Les Perelman’s research on the correlation between length of response and higher scores in SAT writing samples, our scoring tended to rate the longer, typed responses higher. Two instructors wrote their own additional writing prompt for their students.

In a sense, our writing-focused partnership attempted with mixed results to align with what William Condon advocates in “Looking beyond Judging and Ranking: Writing Assessment as Generative Practice.” Condon, whose assessment team designed robust outcomes from a writing test at Washington State University in 2007, argues that when assessments produce writing that is more open-ended, it makes room for graduate students and faculty to bring their own “diverse, non-assessment theoretical frameworks and research interests” to the table (143). Similarly, our project utilized both open-ended writing prompts and a diverse team including a university-level administrator charged with assessing institutional learning effectiveness, a faculty member committed to writing’s potential in interdisciplinary studies, and two graduate students with research agendas in public/professional writing and the rhetorics of race, respectively. While the distinctive interests of the various partners were not so compatible as to make for a straightforward assessment process, the differences in our standpoints led to a more comprehensive examination of student learning.

In the first of a two-part analysis, Branson and Sanchez coded the responses individually against TCU’S QEP outcome rubric for “Discuss Critical Questions,” which draws from the AAC&U VALUE Rubric (“Global”). Each student response received four separate scores (on a range from 1, or “benchmark,” 2 and 3, which are “milestones,” to 4, a “capstone”) for the following categories:

- **Topic Identification**: identifying a “creative, focused, and manageable topic that addresses potentially significant yet previously less explored aspects of that topic.”

- **Description of the Global**: connecting the topic’s relevance “to domestic and global communities” in a way that is “described comprehensively.”

- **Existing Knowledge, Research, and/or Views**: synthesizing information “from relevant sources representing various points of view/approaches.”
• **Analysis:** organizing and synthesizing the evidence in order to “reveal insightful patterns, differences, or similarities related to focus.”

After Branson and Sanchez rated each response individually according to these guidelines, they met to go over scores and compare ratings. They discussed discrepant scores of two or more points, and either came to agreement or averaged their scores. Out of the 620 scores calculated individually (155 students, each having their writing piece evaluated for four categories), discrepant scores emerged 18 times.

By and large, scores in each of the four categories across all seven sections were low (see Table 1). The average score for “Topic Identification” was 2.01; for “Description of the Global,” 1.85; for “Existing Knowledge,” 1.78; and for “Analysis,” 1.76. Particularly unexpected was that classes whose members tended to have attended the public lecture (actually, performed readings from some of Ngũgĩ’s narratives) scored only marginally higher than those classes with fewer members attending. We had assumed that the more direct interaction students had with Ngũgĩ, the more learning they would demonstrate in their writing. Also, the highest scores came in the way of Topic Identification, and very few responses received high scores for “Analysis.” For instance, a student in Robbins’s Contact Zones course wrote that participating in class discussion with Ngũgĩ helped him or her “more fully understand” the author’s work and “experience it in a new way.” This response earned a score of 3 for Topic Identification because it identified a “focused” main idea: the activities increased the student’s understanding of the material (a “milestone” score of 3). However, the student received a 1 for analysis. The response identified a significant takeaway from the experience as hoping to travel to different countries in order to “learn about the different cultures of the world” and to “accept each person and his or her culture as they come.” While suggesting the student had been inspired, the response fell short of offering evidence beyond surface-level platitudes like “we are all raised differently” and “we should never impose our beliefs on others purely because we believe we are superior to them.” This received a 1 for analysis because the response fails to synthesize evidence to “reveal insightful patterns, differences, or similarities related to focus” (needed for a score of 4) and instead “lists evidence, but is not organized and/or is unrelated to focus” (for a score of 1).
While we certainly did not expect everyone to demonstrate a "capstone" level of learning, what surprised us was how frequently students reflected on their personal experiences instead of on their interactions with Ngũgĩ and his work. As Candace Spigelman, Kurt Spellmeyer, Peter Elbow, and others have shown, personal writing is and can be intellectually rigorous. However, in inviting students to share personal responses, perhaps we should have focused our prompts less on having students characterize their own emotional responses, and more, as Katherine K. Gottschalk writes in her analysis of Cornell’s First-Year Writing Seminar, on having students “locate themselves in the personal truths of others.”

From the perspective of program assessment, rather than indicating the Visiting Scholar had little impact on students’ learning, the results may have signaled that we had picked the wrong focus from the available QEP outcomes, that, as discussed above, we had phrased prompts ineffectively, or some mix of these factors (see Table 2).

### Table 1. Scores for Student Responses Using a QEP Rubric

<table>
<thead>
<tr>
<th>Theme</th>
<th>Handwritten response average score</th>
<th>Typed response average score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic Identification</td>
<td>1.63</td>
<td>2.29</td>
</tr>
<tr>
<td>Description of the Global</td>
<td>1.41</td>
<td>2.18</td>
</tr>
<tr>
<td>Existing Knowledge/Research</td>
<td>1.33</td>
<td>2.11</td>
</tr>
<tr>
<td>Analysis</td>
<td>1.26</td>
<td>2.13</td>
</tr>
</tbody>
</table>

Note: Possible scores ranged from 1 to 4.

### Table 2. Percentage of Reviewer 1 and 2’s Combined Scores for Each Theme (n = 155)

<table>
<thead>
<tr>
<th>Theme</th>
<th>Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Topic Identification</td>
<td>31%</td>
</tr>
<tr>
<td>Description of the Global</td>
<td>43.5%</td>
</tr>
<tr>
<td>Existing Knowledge/Research</td>
<td>47.4%</td>
</tr>
<tr>
<td>Analysis</td>
<td>52%</td>
</tr>
</tbody>
</table>
Additional explanations and implications surface, however, when we turn to writing assessment scholarship. From that vantage point, the low scores could have been related to the inefficacy of timed, or rather, on demand responses. Some composition scholars might also attribute low scores to the inefficacy of rubrics themselves in writing assessment. As early as 1986, Loren Barritt et al.’s study of four thousand holistically graded placement essays at the University of Michigan critiqued the use of “pre-established criteria” to assess writing, which, they argue, “makes reading very difficult for anyone who wants to take seriously her own response as an expert, particularly as an expert teacher, while at the same time trying to reach consensus with other readers” (324). And in 1998, Ginnette Delandshere and Anthony R. Petrosky argued that “the generic nature of scoring rubrics” is problematic since it can lead to assessors scrutinizing only “the most visible or superficial characteristics of the performance while neglecting the specific patterns and connections that give it meaning and integrity” (21). This turn toward focusing assessment on the specific, contextual patterns of meaning in student writing led to Libby Barlow et al. calling in 2007 for a kind of “collaborative assessment” that prioritizes “local institutional values and processes over nationally standardized criteria as a basis for understanding student learning” (45). As noted by Broad, related work by scholars such as Grant Wiggins, Huot, and Elbow demonstrates that a standardized writing rubric “documents only a small fraction of the rhetorical values at work” in a text (Broad, What 12). Thus, in our case, the rubric we used may not have highlighted the nuance and specific granularity of thought-through-language within the students’ responses, however short or underdeveloped those texts may have been due to time constraints.

Barritt et al.’s 1986 study is important because it utilized a method relevant to our initial findings. The three researchers each independently read an essay and then discussed the scores over several meetings to adjudicate where to place the student. They reflect that rather than reaching uniform consensus, analyzing student writing together led to broader discussions about writing, and moreover that their “criteria of judgment were adaptable, shifting as we attempted to meet each student on her own ground.” In spite of setting up systematic criteria in advance to judge each essay, they explain, they nevertheless discovered that “engagement of mind by text has the power to disengage even the tightest of preordained scheme”
One takeaway from their study is that in any sort of writing assessment, context matters, and “the predictive validity of a decision about a writing sample is not assured by its being made reliably” (324). In other words, the complex social milieu of the writing-meeting-reader makes for a messy and dynamic evaluative process. These messy rhetorical values are hard to capture in a rubric: “To require evaluators to pay attention only to pre-established criteria is to imply that this is the only way for fair readers to act,” they write (324). In our case, beyond what the rubric could tell us, we needed a different analysis that accounted for nuance, granularity, messiness, and context.

Thus, we conducted a second analysis of the writing. Instead of applying the rubric, we looked for themes that emerged based on our own expertise as teachers and scholars of writing, while simultaneously attempting to reach consensus. This theory of writing assessment resembles what Broad calls “Dynamic Criteria Mapping” (DCM), a qualitative inquiry based on grounded theory in which findings are uncovered inductively through several rounds of empirically gathered data (“Mapping”). In DCM, participants ask themselves what they value in texts rather than relying on predetermined criteria; DCM participants then conduct local, site-based evaluations of texts in collaboration with others through discussion and negotiation, so that shared evaluative frameworks emerge as a set of rhetorical values more nuanced than traditionally designed rubrics. Broad’s use of grounded theory and the qualitative process of Dynamic Criteria Mapping provided a framework for reexamining our data. Adopting a DCM stance revealed how the rubric we had initially been using did not pick up on learning that we value as teachers.

As in the first analysis using the scoring rubric, for the second, more inductive analysis, Branson and Sanchez first worked independently. This time they coded student responses for repeated patterns of meaning, or rather, as Virginia Braun and Victoria Clark describe it, “identifying, analyzing and reporting patterns (themes) within the data” (79). Then, after labeling each response with a range of themes, they met and compared notes. This was not a rigid, formulaic process but rather an exploratory dissection of the student responses using our own personal interests and expertise to guide us (see Table 3). The result was a collaborative report on the additional dimensions of student learning that complicated the rubric’s measures.
Branson and Sanchez brought with them their own research perspectives, which undoubtedly shaped their interpretations of the undergraduates’ responses. For example, Sanchez’s interest in the project stemmed from his background in racial rhetorics and language diversity, leading him to key in on related features/elements in the student writing. For him, what was particularly noticeable was how TCU’s undergraduates wrote about Ngũgĩ’s relationship with race, culture, language, and identity. One student, for instance, stated, “As a typical westerner, we consider imperialism and the modernization of Africa as a good thing, but it may not be what the native people want, and I think that is the biggest takeaway from the discussion with Ngũgĩ.” In the first reading, this answer was scored as a 2 because the response did not elaborate beyond a comment. However, this sentence does illustrate the student’s exploration of positionality, marking a form of learning not specifically identified in the rubric. Americans often consider “modernization” as a marker of success and culture, but in many countries, native people actually resist change in favor of cultural heritage. Here, Sanchez noted, the student writer was fighting against his or her own perceptions and using self-critical reflection.

Branson, who researches public and professional writing, found similar thematic evidence of significant learning despite the low scores garnered through the rubric. Branson noticed trends where students’ responses reflected on how physically seeing a public figure directly contributed to their understanding and appreciation of the texts. “It was good to hear him [Ngũgĩ] read the tone of the book,” wrote one student who had received all 1s in the first round of analysis: he or she could “see how he meant things to be read” (emphasis ours). For other students, interacting with a public figure allowed them to reflect on authorship: “By meeting [Ngũgĩ] and listening to his answers to our questions,” noted another student, who also received all 1s on the rubric, “I got to visualize him as an author who knows what he’s talking about” (emphasis ours). One student wrote that “Asking [Ngũgĩ] … questions made me have a deeper understanding of his work, as I learned about his ideology and passion for writing.” This response received all 1s when judged through our rubric because of its lack of specific examples and for not connecting the learner’s experiences to global communities. What became clear through inductive analysis, though, is that the students quoted above were emphasizing different kinds of learning gains through
their honest reflections on authorship, engagement, and the affective value of interacting with a public figure.

Thus, in their separate second readings, each of the graduate student researchers noted discursive patterns that were different from the rubric’s emphasis. Synthesizing these new patterns, they identified recurring links between the language students used and types of learning being acquired. This analysis posited four learning outcomes: reflective learning, embodied learning, spatial learning, and voluntary learning. Students who showed reflective learning interrogated their own assumptions; embodied and spatial learning critiqued students’ relationships to geopolitical or social, interpersonal boundaries and their physical sense of being through contact with someone different from themselves; and responses demonstrating voluntary learning referenced additional, self-directed learning students had done, beyond course requirements, promoted by the visit. Overall, in this second round of analysis, these emergent themes revealed students questioning their own prior assumptions and predispositions about language, culture, and privilege, despite the lower scores in the rubric-based analysis.

The differences between these results and the results using the rubric led to several productive meetings where we discussed the implications of our findings. Much like Barritt et al.’s 1986 study, our thematic analysis led us to ask ourselves what we valued in a response, what we valued in writing, and what we thought we were measuring with the rubric (was it writing, learning, or both? And what was the difference between the two?). We began to ask ourselves, as teacher-researchers might when faced with perplexing results in the classroom, whether we had even picked the right outcome or asked the right questions initially. These conversations proved vital because, despite the relatively low scores with the rubric, we could now see that engaging with this visiting scholar made some kind of an impact on many students. These findings suggested that future QEP-funded guests could indeed enhance and extend global learning. Meanwhile, our process of working together revealed even more about the value of collaborative assessment, encouraging team members to reflect on their own epistemological frameworks as they had emerged from the second, inductive round of analysis. Our collaboration, in other words, fostered a deep intellectual exchange, strengthening each round of analysis and opening possibilities for future work.
Table 3. Discursive Breakdown of Learning Experiences in New Inductive Analysis

<table>
<thead>
<tr>
<th>Themes found</th>
<th>Definition of theme</th>
<th>Student text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflective learning</td>
<td>Students acknowledged how Ngũgĩ’s visit challenged/changed their opinions, assumptions, and biases</td>
<td>“[Reading his book] changed my opinion because it gave me [a] different perspective on how things were really happening through Ngũgĩ’s eyes rather than interpreting the book on my own.”</td>
</tr>
<tr>
<td>Embodied learning</td>
<td>Students described how Ngũgĩ’s physical presence impacted their learning</td>
<td>“Having Ngũgĩ actually come to class made the book seem so much more real and exciting. I felt much more connected to the text and I understood more why he wrote the memoir.”</td>
</tr>
<tr>
<td>Spatial learning</td>
<td>Students demonstrated efforts to understand various social boundaries and strategies for crossing them.</td>
<td>“As a child who grew up overseas, and my parents still live overseas, I find getting to . . . learn more about other cultures through reading is very important. It allows people to understand how people live around the world, not just in their own country.”</td>
</tr>
<tr>
<td>Voluntary learning</td>
<td>Students expressed that they completed voluntary research and analysis on Ngũgĩ and his work outside of their classroom activities</td>
<td>“After the Ngũgĩ reading, I attended a multicultural awareness event in the TCU commons and I was able to converse with people from all over the globe.”</td>
</tr>
</tbody>
</table>

**Implications**

We see at least three important takeaways from this formative teacher-research project. First, discovering learning that is emergent or unexpected can be valuable for planning and assessing future instructional approaches linked to both university-level and classroom-based special programs. But this discovery process needs to be iterative. Rubrics may be necessary in certain contexts, but in this case and possibly others, they were not sufficient for uncovering the complexities of student writing. By carrying out two separate rounds of analysis, we were better able to recognize and understand multiple, dynamic, and messier aspects of student learning.
Also, through frequent collaborative discussion, our team identified unanticipated learning. Emergent assessment techniques like ours can help generate a more nuanced picture of learning for a course or program, or even across several years in higher education.

The second takeaway is that thematic frameworks can be valuable for assessment projects not directly related to writing programs or writing pedagogy. One fruitful avenue for linking thematic frameworks to writing studies could be through research on opportunity to learn (OTL). To promote equity and social justice within institutional cultures, Diana C. Pullin recommends “insuring that all students, particularly those most at risk of educational failure, are the beneficiaries of an effective opportunity to learn” (334). Open-ended writing prompts can provide such an opportunity, and inductive writing assessments can then underscore divergent learning avenues and outcomes. Edward H. Haertel et al., in that context, suggest that one of the implications of privileging OTL in writing assessment projects is harnessing “the capacity of local interpreters,” or the teachers, administrators, and students involved in the assessment (10). In our case the local, situated, contextual—or, that is, ecological—elements of student learning became more apparent when we approached the assessment from an OTL standpoint.

Having reached that understanding through our inductive analysis, we now appreciate how increased attention and exposure to principles in assessment scholarship can help teacher-scholars develop more pedagogically sensitive and formative assessments. In subsequent work, we can strive to apply these concepts with a better awareness of design parameters around assessment prompts, particularly in terms of what Norbert Elliot describes as a central aim of assessment: achieving “validity associated with fairness,” so as “to assure equitable treatment of those with diverse gender, race/ethnicity, sexual orientation, culture, language, age, disability, and socioeconomic backgrounds” (680). To Elliott’s list, we might add different instructional contexts as a subgroup to identify, since we realized that differences in the course context of learning came into play in our project’s assessments across individual classroom contexts. We do not think every writing teacher has to become an assessment specialist, but we do advocate drawing upon that field’s research frameworks to guide classroom-level teaching and program-level assessment.
This leads to our third takeaway—that a powerful method for drawing upon such frameworks is through cross-disciplinary and cross-role partnerships. Much like the “Integrated assessment” model advocated by the consortium of engaged universities, Imagining America, our assessment project was designed to “integrate the needs and assets of all stakeholders in a co-creative process.” Each team member brought varying levels of expertise to the project, and individuals enacted differing types of engagement, from micro to macro to mixed. Still, from the three teacher-researchers to the assessment specialist and even to the seven instructors who administered the surveys and planned unique course designs supporting Ngũgĩ’s visit—we all valued collaboration. Our formative assessment project was inherently exploratory, looking for emergent patterns of learning. In an early stage of a QEP, we were asking what learning activities might be productive programmatically, but asking in a way that was inviting enough to get teachers to participate. This partnership building fostered collaborative ecologies of emergent assessment, an authentic assessment landscape consistent with the grounded theory offered by Broad and others, but taking into account the practices of collaboration required to get such endeavors off the ground and then sustain them. For example, it was important to give all instructors autonomy over their own courses, including which readings to assign and how to administer the response prompts.

We might describe our particular collaborative ecologies of emergent assessment as cultivating what William Duffy and John Pell call phased collaboration, or rather, a partnership that is both utilitarian (the partnership gets work accomplished) and conceptual (the collaboration improves the work as a whole). Duffy and Pell write that, epistemologically, this kind of collaboration “can result in texts greater than the sum of their individual parts”; within our essay, accordingly, we hope readers will see “the reflexive work collaborators navigate when communicating with each other about not only what to write, but also how best to write it” (248). We each brought different values and research interests to the table, but also a willingness to let our value systems shift and expand, to have our standpoints change by working with people committed to the core goal of helping students learn. In the end, and perhaps ironically because the assessment revealed such
a complicated learning picture, our project prepared us to do a better job for the next group of students involved in university-wide instructional programming. Whether through our own teaching approaches, the rubrics we use, or the kinds of inductive methods we employ to track student learning in the future, all of us involved in this project are better prepared now to formulate, carry out, and evaluate context-sensitive assessment approaches, in our own classrooms and beyond.

**Appendix A: Instructions for Participating Teachers**

Thank you so much for connecting your students to Ngũgĩ’s campus visit. Since this cluster of curricular and co-curricular activities was funded primarily by the QEP, we would very much appreciate your help in gathering assessment data related to student learning outcomes associated with the Global Learning goals of the program. We will be using this particular collection of student writing—along with other data—to help enhance planning for future Visiting Scholar Initiatives and to contribute to the ongoing assessment of the QEP.

Please provide a full set of copies from all your students who respond to the prompts. The copies you submit should be “clean”—i.e., they should not include your own assessment comments or grading, should you choose to use this writing task for a grade. Of course, you should feel free to use this exercise for an in-class assignment, an extra credit activity, a quiz grade—or simply as a formative assessment to aid your own teaching around topics related to Ngũgĩ’s works. We DO ask that you secure responses from all students in your class(es) where Ngũgĩ’s writing was studied, so we recommend that you devote 10 minutes or so of class time to the response exercise to ensure the highest possible response rate.

We are asking all teachers whose classes connected to Ngũgĩ via readings and/or through a class visit to use the same two questions. You may certainly add an additional question or questions of your own if you like; if you do so, you may opt to share that portion of students’ responses with our team or use those responses only for your own purposes.

Thanks again for participating in the visit itself and in our post-visit assessment. Your involvement will help enhance future, similar opportunities for students, ensuring that we benefit from what is learned this time about organizing humanities public programming and about interdisciplinary study supported through the QEP.

**PLEASE RETURN ALL STUDENT RESPONSES NO LATER THAN APRIL 15.**

If you submit digital copies, send them to [NAME], [NAME] and [NAME]. If you return paper copies, you should place them in [NAME’S] mailbox in the English office with a cover sheet that includes your name and your course name and number.
Appendix B: Representative Quotes from Student Writing

Order of Scores Listed Below: 1) Topic Identification; 2) Description of the Global; 3) Existing Knowledge/Research; 4) Analysis

Low Scores

“I am taking away how language is so important to writing, something I had never considered before.” Score: 1, 1, 1, 1

“I hope this can help broaden my worldview and help me expand my knowledge of other countries and their histories.” Score: 1, 1, 1, 1

“I have learned that you have to always do your best and strive to be true to yourself.” Score: 1, 1, 1, 1

“I have a better understanding of imperialism and the effect it had on the people of those colonies.” Score: 1, 1, 1, 1

“We read his book, In the House of the Interpreter, and it gave me a new perspective on what life is like in Africa, and what it is like to live somewhere during a time of struggle.” Score: 1, 1, 1, 1

High Scores

“As a U.S. citizen, it is difficult to relate to Ngũgĩ because I have never had a village torn down and having to relocate the village. The fact that the people in the village had to rebuild and alter gender roles due to the shortage of men is aweing. After hearing Ngũgĩ’s other stories, I realize that I need to start reading his stories from his perspective rather than mine because of my inability to relate to the circumstances. I believe visiting the reading by Ngũgĩ has made me realize that I need to be more empathetic of individuals and their backgrounds and cultures.” Score: 3, 2, 3, 2

“In referencing his works Ngũgĩ said, “Empowerment itself becomes a form of alienation.” This quote was very impactful to me because I had never thought of empowerment in such a way. Ngũgĩ’s response also directly referenced one of the many themes of his short story “Wedding at the Cross” in which Wariuki comes to alienate himself from his wife through empowerment. I previously didn’t link empowerment and alienation but thought of empowerment as gaining a form of capital or utility for the betterment of a community and/or society.” Score: 3, 3, 3, 3

“I was guilty of assuming that English was the highest form of language (in regards to mass communication) because that’s how I’d always experienced it.” Score: 3, 3, 3, 3

“Through the readings, my own research, and his visit, I gained new perspectives on the background inspiration for his stories. Having a better under-
standing for the historical contexts of some of his works, such as the Mau Mau war of Independence seen in “Wedding at the Cross”. He also indulged on the theme of education and alienation and how the two can relate both positively and negatively. Doing some research before hearing him speak gave me a chance to understand connections he made to his pieces of work, as well as his past.” Score: 3, 3, 3, 3

“I am taking away that the experiences he has had as a Kenyan are very much different than mine as a black American.” Score: 2, 2, 3, 3

Acknowledgments
We would like to thank the editorial team and the anonymous reviewers at College Composition and Communication for their critical and encouraging feedback. This essay is stronger, clearer, and more robust thanks to their tireless efforts. We would also like to thank all faculty, students, and staff who participated in TCU’s Quality Enhancement Plan, especially the Visiting Scholar program described here. And finally, we would like to thank Ngugi Wa Thong’o for engaging with students and faculty during his visit in the spring of 2014. If not for his inspirational literary contributions and graciousness with his time, this essay would not have been possible.

Notes
1. The organization’s website gives this self-description: “AAC&U is the leading national association concerned with the quality, vitality, and public standing of undergraduate liberal education. . . . Founded in 1915, AAC&U now comprises more than 1,300 member institutions—including accredited public and private colleges, community colleges, research universities, and comprehensive universities of every type and size.” Besides promoting educational excellence, AAC&U also “works to reinforce the commitment to liberal education at both the national and the local level and to help individual colleges and universities keep the quality of student learning at the core of their work as they evolve to meet new economic and social challenges.” See https://www.aacu.org/about. One dimension of that effort fosters effective models for assessing student learning, as outlined in this report: “Learning and Assessment: Trends in Undergraduate Education, April 2009, https://www.aacu.org/sites/default/files/LEAP/2009MemberSurvey_Part1.pdf.

2. See also Hamp-Lyons, “Farewell to Holistic Scoring?,” which calls for contextualization in assessment and points to rubrics’ limitations. She particularly salutes Broad, “who does not dismiss rubrics but questions, indeed challenges, the value of almost all of them because of their lack of contextual relevance and failure to grow organically from contexts and purposes” (A2).
Works Cited


Gottschalk, Katherine K. “Writing from Experience: The Evolving Roles of Personal Writing in a Writing in the


Perelman, Les. “Construct Validity, Length, Score, and Time in Holistically Graded Writing Assessments: The Case against Automated Essay Scor-
Tyler S. Branson
Tyler S. Branson is an assistant professor of English and associate director of composition at the University of Toledo, where he researches public/professional writing and writing program administration. His published work has appeared in *WPA Journal* and several edited collections. He is currently working on a book project that analyzes what he calls “problematic partnerships” and their impact on the field of writing studies.

James Chase Sanchez
James Chase Sanchez is an assistant professor of writing at Middlebury College, where he studies cultural rhetorics and writing assessment. He is currently completing a book manuscript that analyzes the rhetorical effects of a white preacher self-immolating to protest racism in Sanchez’s hometown of Grand Saline, Texas, in 2014. He recently produced a documentary based on his manuscript research, titled *Man on Fire*, which will be entering film festivals in 2018. His work has appeared or is forthcoming in *Present Tense, WPA, Pedagogy,* and *Steinbeck Review,* and he is a 2015 winner of the Scholar for the Dream Award for CCCC.

Sarah Ruffing Robbins
Sarah Ruffing Robbins is the Lorraine Sherley Professor of Literature at Texas Christian University. Her most recent monograph is *Learning Legacies: Archive to Action through Women’s Cross-Cultural Teaching* (2017). Previous books with a writing studies focus include *Nellie Arnott’s Writings on Angola, 1905–1913; Writing America; Writing Our Communities;* and *Managing Literacy, Mothering America.*

Catherine M. Wehlburg
Catherine M. Wehlburg is the associate provost for institutional effectiveness at Texas Christian University. She has taught psychology and educational psychology courses for more than a decade, serving as department chair for some of
that time, and then branched into faculty development and assessment. She has served as editor of *To Improve the Academy* and is currently the editor-in-chief for the *New Directions in Teaching and Learning* series. Her books include *Promoting Integrated and Transformative Assessment: A Deeper Focus on Student Learning* and *Meaningful Course Revision: Enhancing Academic Engagement Using Student Learning Data.*