

Student-Led Assessment: A Small Study on Classroom Rubric Development and Peer Grading Practices

Brice Particelli

Pace University

Peer review is a common practice in writing studies. However, while there is considerable research on peer review, pedagogical studies on other forms of student-led assessment strategies are less prevalent. This study investigates the expansion of assessment practices into student-led rubric development and peer grading, focusing on their effect on student understanding of the writing process. Utilizing surveys and classroom observations in two second-year composition courses at a university in New York City, this study investigates student-led assessment strategies as a potent pedagogical tool, adding to literature that explores assessment as an active part of the writing process.

Keywords: assessment, rubric, peer grading, peer review, peer workshop, student-led, student-centered, composition, pedagogy, writing studies, grading, profile, genre

Popularized by practitioners of process writing like Peter Elbow and Pat Belanoff (1989), peer review (i.e., peer response, peer critique, or peer feedback) is a common strategy for offering feedback in the composition classroom. Peer review encourages students to take an active role in assessment and helps students develop their understanding of how a piece of writing is written, revised, and assessed (Dixon & Hawe, 2017). Peer review offers benefits to both writer and reviewer, allowing writers to receive additional feedback on their work while encouraging reviewers to articulate their assessment for the benefit of their peers, making it an important formative assessment tool, particularly when guided by a rubric or other specific writing objectives (Lundstrom & Baker, 2009).

The use of rubrics as part of the writing process has been resisted by composition instructors who see rubrics as summative assessments while students see them as a formative tool that connects writing to specific, stated objectives (Reddy & Andrade, 2010). Rubrics, just like a magazine or journal's submission instructions, offer tools for writers to see what their audience values. Student-led rubric development offers additional opportunities to engage students in (a) process-focused discussions on the connection between writing and writing assessment (Huot, 2002b; Sadler, 1998) and (b) critical discussion on why communities value certain forms of language over others and how equity and exclusion factor into writing and writing assessment (Inoue, 2015).

This research expands the literature on peer review and rubric use in the classroom, exploring the use of rubrics in additional aspects of the writing process. This study examines surveys and classroom observations from two second-year composition courses conducted by the researcher. Students in these courses worked together to develop a rubric for an assignment and then used it for peer review and peer grading. Uncovering the potential for these student-led assessment strategies, the findings from this study support the idea that professors empower students by having them develop and implement rubrics, indicating that these practices are a potent pedagogical option.

This article begins with a review of the literature surrounding student-led assessment strategies, followed by the methodology for this small study. The article then offers an overview of the course

and classroom practices during rubric development and peer grading, followed by the study results, conclusions, and considerations for further study.

My aim in initiating this study is to consider the uses of peer-led assessment strategies as a formative and process-based pedagogical tool, involving students in as many aspects of the assessment process as possible to encourage them to better understand the connection between writing and assessment. Some guiding questions included: Do students find peer assessment valuable? In what ways do students learn from engaging in the process of peer assessment? How is a student's confidence as a writer affected through the process? What implications could the study's results have for the way instructors incorporate assessment in the classroom?

Literature Review: Student-Led Assessment Practices

Instructor resistance to using rubrics is rooted partly in the history of rubrics, which, as Broad (2003) wrote, runs parallel to the history of standardized testing. Rubrics emerged from a long line of standardized tests born during World War II when the government began offering an ever-changing set of desired learning outcomes used to establish standards that often culminate in high stakes, summative assessments and placement exams (Broad, 2003, p. 4). This kind of institutionalized, top-down rubric design was viewed by Broad as reductive, and other scholars have criticized these rubrics as ignoring process-oriented pedagogy and classroom-based practices (Heritage, 2010; Martins, 2008). Broad argued that rather than offering rubrics to highlight the complexity and context-sensitivity of rhetorical situations and assessment, "we have presented our students with a process and document born long ago of a very different need: to make assessment quick, simple, and agreeable" (p. 4).

While instructor-developed rubrics need not fall into this simplistic, top-down trap, writing instructors still view rubrics with skepticism. In a 2010 review of rubric use in higher education, Reddy and Andrade found that most college writing instructors are resistant to the use of rubrics, often seeing them as summative assessments meant to standardize grading protocols rather than as teaching tools. Students, on the other hand, see rubrics as helpful "because they clarify the targets for their work, allow

them to regulate their progress and make grades or marks transparent and fair” (p. 438), particularly when those rubrics are offered early in the writing process (p. 439). This difference between student and instructor perception of rubrics is striking. Reddy and Andrade concluded that instructor resistance can be problematic because research shows that the result of using rubrics is “higher achievement and deeper learning” (p. 439).

Sadler (1998) argued that students should not only have access to rubrics but also be involved in writing rubrics. Rather than only being able to recognize and solve externally sourced problems, students can then learn how to frame problems “as part of their progressive journey into self-assessment, and at more advanced levels, as a key skill for professional life” (p. 81). Helping students frame their own learning allows students to embrace the authority inherent in assessment (Huot, 2002b) and “recognize how ubiquitous [assessment] is within the process of reading and writing” (Huot, 2002a, p. 4). Similarly, Inoue (2005) suggested that aside from becoming more active learners, students also become more self-conscious and reflective writers if they are involved in developing rubrics (p. 209). While the time it takes for rubric development can feel “chaotic and unproductive,” Inoue wrote, “the point is to have the discussion and begin to cultivate a culture of dialectical vying” in order to “problematize [students’] notions of some static, essential, ‘correct’ assessment or grade that goes with each piece of writing” (p. 216). This must include rubric development as well as rubric use, concluded Inoue.

The empirical research on peer grading in college writing courses is largely limited to the use of instructor-created rubrics within online learning and is focused on validity (students’ ability to assess each other similarly to the instructor’s assessment) and reliability (students agreeing with each other on assessment) as well as student and instructor perception of peer grading. In a 2006 study of 708 online writing students at four universities, Cho, Schunn, and Wilson found that when at least four students grade a piece of writing, instructors reported that the peer grading was as valid and reliable as the instructor’s own assessment in 95% of the student grades. In a 2011 study of 250 students across six universities who were using an online peer assessment system in writing across the disciplines courses, Kaufman and Schunn found that students initially

perceived the fairness and validity of peer grading as low. However, perceptions increased significantly after participation. The researchers also found that students reported having a more positive experience when peer feedback and grades were accompanied by instructor feedback and grades (as opposed to peer-only assessment).

These studies on peer grading (a) reveal students' ability to implement grading protocols and (b) offer the potential of extending community-based rubric development into peer grading while pointing to potential downfalls. Many studies (Kaufman & Schunn, 2011; Liu & Carless, 2006; Smith, Cooper, & Lancaster, 2002) found that discomfort with peer grading is linked to students' lack of trust in the expertise of their peers and the unfamiliar power dynamic of peer assessment, but the studies also showed that offering students training and guidance in the use of the rubric before they used it to grade each other helped mitigate those fears.

While there is a wealth of research regarding the use of instructor-developed rubrics that involve online composition students in peer grading (Cho et al., 2006; Kaufman & Schunn, 2011; Liu & Carless, 2006) and much discussion by composition instructors on the benefits of student-led rubric development (Huot, 2002b; Inoue, 2005; Reddy & Andrade, 2010; Sadler, 1998), there is a lack of empirical research regarding the practice of cocreating rubrics with composition students. This study explores that gap, focusing on student perception of peer-led assessment practices that extend from rubric development to peer grading.

Methodology

This IRB-approved research project focuses on two sections of a course taught by the researcher titled "Writing in the Disciplines," a required second-year writing course at a small university in New York City. The undergraduate population of the university self-identifies as 49% White/non-Hispanic, 14% Hispanic, 12% Other, 10% African American, 8% Asian/Pacific Islander, 4% Multi, and 3% Unknown, as well as 59% female and 41% male (Pace University, 2018). We did not collect demographic data as part of this study, but the two class sections roughly reflected these demographics. Both sections ran Tuesday/Thursday for 15 weeks

(30 sessions), 1 hour and 25 minutes each session. One section had 23 enrolled students, while the other had 20. All students consented to participate, allowing for a study size (N) of 43 students.

This study investigates how students perceived their learning during student-led assessment within a high-stake writing assignment (20% of the course grade). We implemented community-based rubric development in each classroom, and students played a central role in determining their peers' grade. Surveys were taken at three points during the semester:

- An “initial survey” in the first week of the 7-week assignment (Appendix A);
- A “pregrading survey” at the beginning of the seventh week, 2 days before students graded each other (Appendix B); and
- A “postgrading survey” 7 days after grading (Appendix C).

The surveys focused on how students perceived their classroom assessments and learning, specifically regarding the use of rubrics, peer review, peer grading, and instructor grading. The surveys were developed using language found within the literature of peer assessment practices. Surveys were optional, anonymous, completed in class, and comprised mostly of Likert-type scale statements, along with a few rank order and open-ended questions. They were designed and collected in Qualtrics.

Rather than try to minimize the research bias of the Hawthorne Effect, or “observer effect,” where subjects modify their behavior because they know their actions are being observed (Monahan & Fisher, 2010), the surveys were integrated into the pedagogy. After each survey, the class discussed the results, leading to conversations on the ways in which students understood how peer feedback, rubrics, and assessment fit within writing practices across disciplines. This use of surveys (to create a conversation) embraces the observer effect, allowing research subjects to know they are being observed in order to encourage the subjects to think more critically and openly (Monahan & Fisher, 2010, p. 358). In fact, using the observer effect as a teaching strategy—a regular classroom practice—launched discussions on assessment, genre, and the social context of writing while encouraging students to reevaluate and reflect on the broader contexts of peer-, self-, and instructor-led assessment. It should be kept in mind,

then, that this study's results might be dependent on the use of surveys, or a similarly structured metacognitive "pause," to deepen discussions by regularly connecting a classroom's assessment practices with the way that writing is assessed across disciplines.

This study focused on a convenience sample of 43 students in two classes, asking students to participate anonymously in three surveys throughout the project. During each of the three surveys, students received an emailed link at the beginning of the session and were offered time to complete the survey on their phone or laptop. Students who did not attend class that day were then also able to take the survey.

At the conclusion of the semester, the survey results from both classes were downloaded. Analysis of the results across the three surveys focused on how students perceived the assessment process at three stages: before the project began, after rubric development but before peer grading, and after students received their grades. As we analyzed the responses, we focused on how students' perceptions changed throughout rubric development and peer grading, as well as what students perceived they learned from the process.

Context: Classroom and Assessment Practices

"Writing in the Disciplines" is described in the course catalog as focusing "on writing effective essays and research papers in disciplinary modes and in students' field of interest. It may include interviews, analysis of journal articles, and appropriate documentation style formats" (Pace University, 2019). This is, of course, a broad mandate. Each classroom might have art students alongside accounting students, which could pose a problem when discussing how writing functions differently from one student's fields of interest to another's, especially when each discipline's writing conventions are often significantly different.

Rather than seeing this as a problem, our department leadership sees this as an opportunity to design student-centered courses, encouraging instructors to pursue student-led investigations of language, research, and writing within their own fields. While specific mandates, training, and curriculum requirements are minimal, recent professional development meetings for the composition faculty have been centered on genre studies,

writing about writing, writing for publication, peer review, plagiarism, and research practices.

My approach to the course was built from my own experience and training. During my MFA in writing, my most formative classroom experiences were in “form and technique” courses, in which we dissected contemporary literature with an eye toward writing our own work. During my PhD in English education, I ran a writing- and publishing-focused nonprofit working with underserved New York City high schools where we asked students to investigate a genre together and then to write for publication, culminating in a glossy, ISBN-numbered publication that students would take home for their bookshelves. Finally, since beginning teaching at my current university 5 years ago, discussions with fellow composition faculty members have furthered my interest in rhetorical genre study’s place within the approaches laid out in Wardle and Downs’s *Writing About Writing* (2017), which focuses on unveiling and articulating the writing process through research, “introducing students directly to what writing researchers have learned about writing and challenging them to respond by writing and doing research of their own” (p. v). Through these and other experiences, I have developed my approach to this course to help students investigate, research, and discuss the ways writers use and adapt genre conventions to engage with disciplinary modes and discourse communities. We begin with foundational discussions, and then we engage in two major assignments: a profile of an individual in each student’s own field followed by a genre analysis of a form of communication within each student’s field.

During this study, I began my two sections of the course with some baseline readings, including an excerpt from John Swales’s “The Concept of Discourse Community” (2017). This not only allowed students to be introduced to terms like *discourse community*, *genre*, *discipline*, *conventions*, and *performance* in the context of composition studies, but also allowed students to begin questioning the “why” of writing rather than focusing on the “how.” Our first week included discussions about students’ past writing assignments, and in most cases, students reported that academic assignments were assessed by an audience of one (their instructor) and were rarely tied to a specific, publicly used genre.

Beginning the course with these discussions allowed students to question their own understanding of writing and better understand how they can accommodate to a specific audience's expectations (their rubric), whether that audience is one or a thousand. "Lightbulb moments" in those discussions came when students realized why they felt their writing was "B.S." sometimes, even when it was perceived as "good" by their instructors—which often occurred when students said their writing was done to accommodate a single-person audience in language that felt forced and false. It helped facilitate the future discussions we would have as a class when we read exemplar texts, asking students to consider the link between genre and rhetorical situation as we read profiles written for different purposes and audiences.

Our discussions in that first week freed students from prior misconceptions and allowed them to see that each new writing assignment or new genre requires understanding the conventions and assessment practices of a specific community. These discussions defined our classroom culture as one in which we would question the social context of writing and discuss the fluid nature of genre and language rather than focus on learning rules, templates, or fixed genres.

During the second week of the course, we began a genre study of the "profile essay"—a journalistic genre in which writers choose a person, research the person and their field (or context), interview the person, and then write an essay that fits within this broadly defined but fairly common form. Focusing on this flexible genre allowed us to explore exemplar texts across form, style, voice, and audience and discuss the socially situated nature of writing. Because these texts came from a broad selection of publications, some texts utilized slang and dialect, some texts were in a traditional journalistic voice, and some texts were multimodal. This diversity in texts encouraged discussion on the role of translingual and diverse composing practices essential to genre studies (Gonzales, 2015), allowing students to consider their own writing practices within this and other genres. Similarly, the genre's flexibility enabled research-focused students to dive deeply into research and narrative-focused students to dive into observation and interview while still requiring all students to engage in a bit of everything. Finally, the profile's focus on how an individual acts

within and represents a field also offered parallels to the assignment they would embark on after this one—a genre analysis of a communication practice within their discipline—which allowed for strong recursive opportunities throughout the course.

We began the 7-week assignment as readers, analyzing diverse examples from the profile genre in increasingly complex analyses conducted through discussions and weekly mini-essays. Analyses and reflections were slowly replaced by drafting as students attempted techniques seen in the readings. We also established a few regular classroom practices, including

- reading one to two exemplar texts per session for 5 weeks, followed by required but self-directed readings in which students chose exemplar readings that best suited their own projects;
- focusing a discussion on one to two aspects of form, technique, or context regarding the day's readings;
- following most discussions with an in-class writing exercise that emulated a specific approach we had just discussed;
- using our online discussion forum to write a 400-word analysis of the readings each week for 4 weeks, followed by drafts of sections of students' essays for the remaining 3 weeks (based on emulation);
- reading profile essay definitions, submission guidelines, and commentary by a profile writer each week to discuss how genre and assessment relate to each other and to consider context and approach;
- regularly discussing how the profile genre's approaches to research, style, structure, and so forth relate to students' own field's genres; and
- taking three surveys—two during the assignment and one after—that opened discussions on how writing is assessed across disciplines and rhetorical situations.

The short-term goal was to help students understand the movable borders of this genre so they could understand their options for this assignment and their audience (their peers). The long-term goal was to help students view genre as a communication structure that adapts to fit a community's needs and expectations and to help students understand that all writing is assessed through context-based criteria and socially-situated expectations, whether a grade is attached to the writing

or not. These practices allowed for a genre study from which we could develop a rubric to fit our own situation and then use that rubric to have students grade each other. As Sadler (1998) stated, it helped students “frame problems themselves” as part of their “progressive journey into self-assessment” (p. 81).

By the time the students’ first full draft was due, we had read approximately 20 exemplar texts, discussed repeatedly why forms and language look different in publications, and read several definitions of “profile essay,” including the two below:

A “profile feature” is a newspaper article that explores the background and character of a particular person (or group). The focus should be on a news angle or a single aspect of the subject’s personal or professional life. (“How to Write a Profile,” 1999)

A Profile is a biographical piece—a concise rendering of a life through anecdote, incident, interview, and description (or some ineffable combination thereof). (Rothman, 2012)

By the time we got to our own definition of “profile essay,” we were ready to develop a rubric that would fit the work students wanted to do. I imposed a genre category and a suggested word count (2,000–3,000 words). Students then had to determine broad enough criteria for our rubric without making the criteria so broad as to lose the genre or so narrow as to eliminate options for the writer’s choice.

While we had discussed the genre each week (assessment conversations were ever-present), the rubric was not developed until weeks 4 and 5. It was created during an hour of one session and 15 minutes of two other sessions. In our first rubric creation discussion, I asked students to create their own detailed definition of the profile genre using a technique called “snowballing.” I have heard instructors refer to “snowballing” in teacher workshops but have not seen references to it in literature. To “snowball,” an instructor asks students to write their own ideas on a piece of paper—in this case, listing the genre’s essential elements. Students are then instructed to get into pairs, sharing their ideas with a partner and negotiating a new list. Finally, I handed out sample rubrics (of other genres) and blank rubric tables and asked the students to get into groups of four

to negotiate a new list and fill out the rubric elements (e.g., the rows) as well as descriptions of what an “A” would be for each of those elements (e.g., the first column). When they were done, the five groups of four each wrote their proposed rubric elements on the board, and then the whole class discussed and debated our options. While I facilitated and structured discussions, I tried to keep my influence to a minimum, and the only direct influence I exerted was to push back against each section’s desire to have a row devoted entirely to “grammar.” In both class sections, I encouraged the students to broaden the category of “grammar” in order to push students to focus on ideas as essential elements and avoid an over-emphasis on Standard American English. This moment, like many others within rubric development, became part of our ongoing discussions about how rubrics can silence some voices by demanding specific dialect and language use and encouraged students to consider what was essential versus what was academically expected.

The elements developed by each section were:

Section 1

- A biography or history of the subject
- Quotes from and observations of the subject
- A “broader perspective” or angle on the subject(s)’s place within a discipline, field, or idea
- Organization, clarity, and grammar

Section 2

- Background: biography of subject and/or a history of the field
- Interview and/or observation: interactions with subject(s)
- Author’s voice or presence works well with profile subject
- Subject’s impact or relevance (or place within) a field
- Grammar, structure, and other style and clarity issues

I then collected the students’ completed rubric tables with their descriptions for what an “A” would be for their proposed elements, letting them know that I would synthesize their language as I fleshed out each rubric in greater detail into our proposed rubric. For the following two sessions, we applied this proposed rubric to professionally written exemplar texts and student-written exemplar texts and then discussed and

revised our drafted rubric for approximately 15 minutes until we came to the final rubric (see Appendix D for Section 1's rubric).

These stages of the development and application of the rubric acted as the beginning of rater training, in which students established an understanding of our rubric and prepared themselves for peer grading. While typical rater training asks raters to read sample papers that are representative of scores within a rubric and calibrate their scores as a group before they score independently (Huot 2002a, pp. 85–86), our rater-training process was more limited. We applied the rubric to two student-written profiles from past classes and two professionally published profiles. We then applied the rubric to peer drafts as part of a peer review day and finally to students' own drafts for self-assessment.

During week 7, the students graded each other. Each student uploaded one copy of their final profile essay for me and brought two copies to class for their peers to grade. I collected the papers and redistributed them with blank rubrics that included a space for written comments. I then asked graders to place their name on a post-it, which I would remove. While peers saw the name of the person they were grading, their own grading was anonymous. While I did not grade the grader, I wanted a level of accountability in asking students to claim their work as graders.

As a final way to alleviate pressure, I reminded students that I allow for rewrites. While it was on the syllabus, and discussed during the first week, this last-minute reminder freed students to approach their comments (and assessment in general) as formative, not summative. While few students ever take me up on rewrites (in this case, only four of 43 did), it allowed the student graders to shed the fear of issuing "final" grades to their peers. I also reminded the students that I would be grading the papers independently, and that while I would defer to the average of the two student grades, I would add my grade into the average if the grades were more than 2% away from my own independent grades, and I would override that grade entirely with my own if the student-given grades were still significantly lower.

Each student then graded two of their peers' papers during a class session. Students were free to leave once they had graded both papers, but I did not tell them that until the first students had finished. Most students took an hour to read and grade both essays.

I graded all essays independently, using their online submissions. I used the same rubric the students used, and I graded the submissions before I looked at peer grades. When I returned the papers with final grades, I included the two peer-graded papers with filled-out rubrics along with a print-out of my written comments. The two peer grades and a final grade appeared at the bottom. I did not include my rubric sheet or grade, only the two peer grades and the final grade, and I described my approach to averaging the grades as I handed back their work. None of the students challenged their grade.

Results

I conducted three anonymous surveys in each class section throughout the 14-session, 7-week assignment. The initial survey focused on assessment in general (Appendix A), while the pregrading survey (Appendix B) and postgrading survey (Appendix C) focused more on peer grading. Of the 43 students, 38 took the initial survey, 39 took the pregrading survey, and 37 took the postgrading survey, indicating that some students were absent or chose not to participate in the optional surveys.

In the initial survey, I was interested in how students perceived peer review, rubric use, and grading. Sixty-three percent of students agreed or strongly agreed that “professor-directed peer feedback on written assignments has been valuable in my past courses,” which is in line with other findings that suggest that structured peer review is valuable (Lundstrom & Baker, 2009).

However, students perceived their instructor’s role in formal assessments as more nuanced. While 59% of students agreed or strongly agreed that in their college courses, “requirements and grades tend to be clear” and 67% agreed or strongly agreed that “exams and written assignments tend to be fair and reflect course content,” there were indications that the clarity in these courses did not always come from rubrics or well-articulated assignments. To the statement, “In my experience, professors use rubrics (or well-articulated assignments) to assess written assignments fairly,” 51% were neutral, 31% agreed or strongly agreed, and 18% disagreed or strongly disagreed. It appears that students gain clarity in part from their fellow students since 68% agreed or strongly agreed to the statement, “Working and ‘talking it through’ with fellow students has

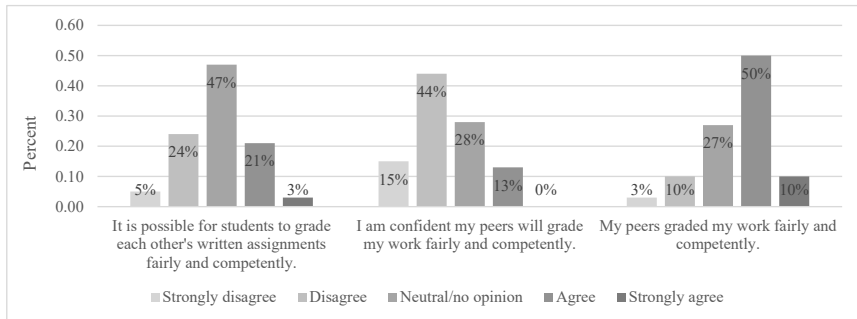
been essential to understanding a professor's expectations on written assignments." In classroom discussions, students expounded that they typically receive a difficult-to-understand assignment sheet without a rubric and that the professor's expectations become clearer through classroom and peer-to-peer discussions, which is in line with the finding that instructors tend to avoid rubrics or otherwise clearly defined assignments (Reddy & Andrade, 2010).

Some of the most interesting findings came from how students perceived peer grading. A series of questions repeated in all three surveys centered on how students perceived their ability to grade each other's work "fairly and competently" (see Table 1; Figure 1). In the initial survey, 47% of students were neutral when asked if peer grading could be conducted fairly and competently, and 29% disagreed or strongly disagreed. In a classroom discussion, students described the same reticence that is reflected in the literature, in which students do not trust each other's expertise and are uncomfortable with the power dynamic (Kaufman & Schunn, 2011; Liu & Carless, 2006). One student stated, "I just don't want another student to have that kind of power over my life," emphasizing the word "life" with worry. Students were unsure if their peers had the expertise to grade their work. Then, despite a significant amount of time spent on rubric development and assessment discussions, their fear spiked in the pregrading survey. Two days before peer grading, 59% of students disagreed or strongly disagreed that their peers could grade each other's work fairly and competently, a significant increase from the initial 29%. However, in the postgrading survey that number reversed again. Only 13% disagreed or strongly disagreed that their peers had graded fairly, while 60% of students found their peers to be fair and competent graders. (The median and mode reflect those numbers as well, while the mean flattens out for the first two surveys but still reflects a spike from the postgrading survey responses.)

Table 1
Student Confidence in Peer Grading

| Survey | <i>n</i> | Survey prompt | 1 Strongly disagree | 2 Disagree | 3 Neutral/ no opinion | 4 Agree | 5 Strongly agree | Mean | Median | Mode |
|-------------|----------|---|---------------------------|---------------|-----------------------------|------------|------------------------|------|--------|------|
| Initial | 38 | It is possible for students to grade each other's written assignments fairly and competently. | 5% | 24% | 47% | 21% | 3% | 2.92 | 3 | 3 |
| Pregrading | 39 | I am confident my peers will grade my work fairly and competently. | 15% | 44% | 28% | 13% | 0% | 2.85 | 2 | 2 |
| Postgrading | 37 | My peers graded my work fairly and competently. | 3% | 10% | 27% | 50% | 10% | 3.70 | 4 | 4 |

Particelli, B. (2020). Student-led assessment: A small study on classroom rubric development and peer grading practices. *Journal of Response to Writing*, 6(1), 42–75.

Figure 1*Student Confidence in Peer Grading Across Three Surveys*

While the spike in student concern over the fairness and competence of their peers right before the peer grading exercise might indicate a fear of a lack of preparation, other results indicate otherwise. While students were wary of their peers just before grading, with only 13% agreeing that their peers could be fair and competent, they nevertheless expressed confidence in their own ability to grade. In the pregrading survey, 54% of students agreed or strongly agreed to the statement, “I understand the genre well enough now to grade fellow students’ work competently,” with 36% neutral and 10% disagreeing or strongly disagreeing. This indicates that students were comfortable with the assignment and the rubric, but uncomfortable with the situation.

Regarding the consistency of final grades, resulting grades indicated strong community-wide understanding of our standards. The average grade offered by the two peer graders (87.4%) was surprisingly close to my own independent blind grading (87.9%). However, individual grades indicated more variance. While 60% of individual student grades were within 2% of my own, 20% were 5% different or more. In practice, I was able to side with the average of the two student grades in all but five cases, but if the goal is reliability—to replace instructor grading with student grading—five in 43 cases is not ideal, which might support Cho et al.’s (2006) findings that a minimum of four student graders is necessary.

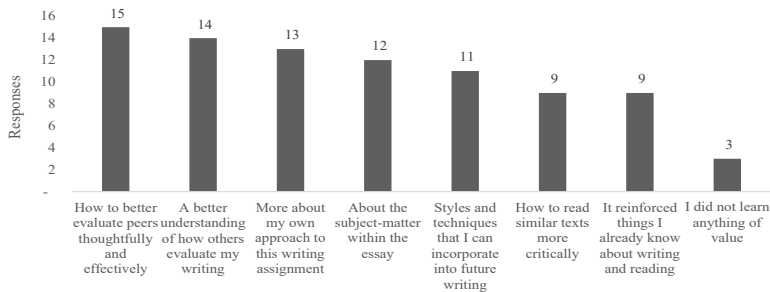
One of the most surprising results was how students perceived their understanding of the genre and assignment. Eighty-nine percent agreed or strongly agreed that “I understand the basic elements of this genre,”

and 97% agreed or strongly agreed that “I understand the requirements of this writing assignment.” For many of the instructors whom I have discussed these findings with, as well as in my own experience, this level of understanding of an assignment is significant. It indicates that the process of genre study, rubric development, and rater training allowed students to become deeply familiar with the genre, their audience, and audience expectations, likely creating a situation in which grades were higher because expectations were clear. Equally striking, and perhaps related, in response to the pregrading survey prompt, “I am proud of the work I have been doing on this piece of writing,” 75% of students agreed or strongly agreed, while 23% were neutral. Only one student disagreed.

A central aspect of this study was to measure how students perceived their own learning throughout the process. In response to the postgrading survey statement “Grading my peers’ writing, I learned (check all that apply),” the students most often chose that they had learned “how to better evaluate peers thoughtfully and effectively,” “a better understanding of how others evaluate my writing,” and “about my own approach to this writing assignment” (Figure 2). Unsurprisingly, by applying their community-developed rubric to peers’ work, students learned how their

Figure 2

Student Perceptions of Peer-Grading Learning Outcomes



Note. Student perception ($n = 38$) of peer-grading learning outcomes in regard to the postgrading survey statement “Grading my peers’ writing, I learned (check all that apply).”

own and others' writing is evaluated and assessed. Perhaps most promising, though, these results indicate that students saw their extensive involvement in assessment as a skill that they could use in future writing tasks—the top three responses focused on transferrable skills.

Finally, when asked in the postgrading survey if this peer assessment strategy should be repeated in future courses, 57% responded “yes,” 23% responded “yes, with some changes,” and 20% responded, “no.” Thus, 80% of students saw the benefit of incorporating peer grading into a required writing course. Classroom discussions revealed that remaining hesitations centered on individual graders being too harsh, a desire for graders to be graded, and concerns on student expertise as graders. However, there was also an overwhelming student desire for more instructors to incorporate some version of community-centered rubric development as a common classroom strategy. These findings offer strong initial support for that desire.

Conclusion and Suggestions for Further Research

Grading is sometimes seen as an impediment to writing because it draws students' attention away from more formative and qualitative feedback (Heritage, 2010; Martins, 2008). It is also problematic, as Brookhart's (1994) often-cited study concluded that instructors' grades are a “hodgepodge grade of attitude, effort, and achievement” (p. 279), made even more worthy of scrutiny among discussions of the role of bias in assessment (Inoue, 2015).

This study takes up the issue of how one might localize writing assessment through a dialogic engagement that involves students in genre study, rubric development, peer review, and peer grading. The aim of this study was to consider how assessment could be used more effectively as a process-based pedagogical tool, involving students in as many aspects of the assessment process as possible to embolden students to better understand the connection between writing and assessment.

Reflecting the literature, this study's results show that through involvement in assessment practices, students exhibited a clear understanding of the genre they were developing the rubric for, indicating as Sadler (1998) wrote, they felt able to recognize, solve, and frame problems themselves.

After developing and using rubrics, students could also better evaluate others and understand how others evaluate them (Reddy & Andrade, 2010; Sadler, 1998), suggesting opportunities for the transfer of these skills across disciplines. While rubric development took time, students indicated that “dialectical vying” was worth taking the time to challenge and problematize students’ notions of a singular version of a “correct” assignment (Inoue, 2005) despite discomfort with an unfamiliar power dynamic in the days before peer grading (Kaufman & Schunn, 2011; Liu & Carless, 2006; Smith et al., 2002). This study extends online peer grading research into an in-person course with a student-developed rubric and also supports the literature that encourages involving students in developing and using rubrics. This support was perhaps most striking in the following two findings: (a) 97% of participants said they understood the requirements of the writing assignment, and (b) students reported their top learning outcomes from the study as the ability to evaluate others’ writing and to understand how others evaluate their writing, two outcomes that indicate significant opportunities for transfer (Figure 2).

While this study indicates that students felt that they were empowered by their participation in the assessment process and that they had learned about the relationship between assessment and writing, questions arose that need further study. Our rubric development might have relied too heavily on in-class participation. While the use of “snowballing” allowed individuals to articulate their ideas before we turned to group consensus, as the group got larger, the final decisions were ultimately made by the most vocal students. This is a potential problem for second-language learners and second-dialect speaking students, as well as introverted students who might feel silenced by the decision-making process. Including an online aspect to the process would offer additional openings for students who are more reticent public debaters. Rubric development was also heavily guided by the instructor. Debates on revisions were moderated and translated into the rubric by the instructor rather than the students, and the actual writing of the rubric, while based on student-led choices, was again done by the instructor rather than the students. These actions likely altered the final rubric to some degree, silencing some student voices.

The biggest challenges remain in grading protocols. As Kaufman and Schunn (2011) found with their online writing students, students in this study perceived the fairness and usefulness of peer grading lower right before grading and significantly higher after participation. However, this study found a more nuanced response when discussing the process with students in the classroom afterwards. While the survey results supported Kaufman and Schunn's findings, a small handful of students described being upset or disappointed by their fellow classmates' assessment when they were given two very different grades. Instructors might benefit from following Cho et al.'s (2006) findings, which suggest that having four graders offers greater reliability and offers students access to more peer feedback. One might also consider grading the graders on the quality of their written feedback to encourage a deeper critical engagement with the assessment process and to reduce concerns of social-based grading, in which students often know whose work they are grading in an in-class setting. Alternatively, an instructor with two sections could have one class grade the other, which would have the added pedagogical benefit of discussing each section's rubric, but the drawback of this is asking graders to utilize a rubric they did not create.

However, any hiccups in grading protocols also allowed for rich classroom discussions on grading inequities in general, the different expectations and backgrounds of assessors (including instructors), and our expectations regarding language use and language difference. These moments offered opportunities of learning for the classroom as a whole. An expanded study might therefore benefit from including follow-up interviews with students on whether or not they internalized these lessons and applied them to future writing projects. A longitudinal study could allow an in-depth study that includes demographic data.

Regarding the overarching practice of student-led rubric development and peer grading, it is worth questioning whether future student perceptions of their learning outcomes would be similar to the ones in this study's findings if only the instructor did the grading while students simply focused on rubric development. There are benefits and drawbacks to removing students from the grading process. It is true that the majority

of the students' time spent on assessment was centered on developing, discussing, and calibrating both the rubric and the writing to fit the community's expectations. Grading was merely the last step. However, grading is a meaningful action in our current institutional structure, and to involve students as graders is to empower them. Removing grading would minimize some of the consequences and benefits of rubric development. Alternatively, Inoue (2005) offers an interesting option for rubric use in which students codevelop a rubric and then the instructor and students individually meet to determine the grade, coming "to some agreements about their portfolio grade in private conferences that they manage" (p. 211). Of course, any involvement of students in the grading process will always be fraught with issues that step beyond the classroom when we work within institutions that use grades to determine financial support, fellowships, acceptance into graduate school, and employment opportunities.

We need additional research that focuses on how student-led rubric development and use affects student learning. Involving students in the assessment process helps them develop the skills needed to frame a task themselves, write toward specific goals, and then assess writing based on criteria that exists explicitly within a community. While peer grading research has focused on reliability and efficiency in online courses, additional research is needed that focuses on pedagogical concerns. We need to find a way to help students see assessment as interwoven within the writing process, the connective tissue between text, writer, and audience.

Involving students in genre study, rubric development, peer review, peer grading, and the conversations that are required to engage in these steps allows students to question the process of assessment from the perspective of both writer and reader, assessed and assessor, and to consider how to apply those skills beyond the classroom. The results of this study indicate that student-led rubric development and peer grading leads to a better understanding of self-evaluation, peer evaluation, and students' own writing processes. Findings also indicate that students consequently have greater pride in their work and a deeper understanding of the assignment and genre.

Further pedagogy-focused studies in peer-led rubric development and peer grading is encouraged. Understanding the efficacy of these practices will be essential for further discussion on how we might continue connecting assessment within discussions of the writing process.

References

- Broad, B. (2003). *What we really value: Beyond rubrics in teaching and assessing writing*. Boulder, CO: University Press of Colorado.
- Brookhart, S. M. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education*, 7(4), 279–301.
- Cho, K., Schunn, C. D., & Wilson, R. W. (2006). Validity and reliability of scaffolded peer assessment of writing from instructor and student perspectives. *Journal of Educational Psychology*, 98(4), 891–901.
- Dixon, H., & Hawe, E. (2017). Creating the climate and space for peer review within the writing classroom. *Journal of Response to Writing*, 3(1), 6–30.
- Elbow, P., & Belanoff, P. (1989). *A community of writers: A workshop course in writing*. New York, NY: Random House.
- Gonzales, L. (2015). Multimodality, translanguaging, and rhetorical genre studies. *Composition Forum*, 31. Retrieved from compositionforum.com
- Heritage, M. (2010). *Formative assessment and next-generation assessment systems: Are we losing an opportunity?* Washington, DC: Council of Chief State School Officers. Retrieved from <https://files.eric.ed.gov/fulltext/ED543063.pdf>
- How to write a profile feature article. (1999). *New York Times: The Learning Network*. Retrieved from <https://web.archive.org/web/20180117055615/http://www.nytimes.com/learning/students/writing/voices.html>
- Huot, B. (2002a). *(Re)Articulating writing assessment for teaching and learning*. Logan, UT: Utah State University Press.
- Huot, B. (2002b). Toward a new discourse of assessment for the college writing classroom. *College English*, 65(2), 163–180.
- Inoue, A. B. (2005). Community-based assessment pedagogy. *Assessing Writing*, 9(3), 208–238.
- Particelli, B. (2020). Student-led assessment: A small study on classroom rubric development and peer grading practices. *Journal of Response to Writing*, 6(1), 42–75.

- Inoue, A. B. (2015). *Antiracist writing assessment ecologies: Teaching and assessing writing for a socially just future*. Fort Collins, CO: WAC Clearinghouse.
- Kaufman, J. H., & Schunn, C. D. (2011). Students' perceptions about peer assessment for writing: Their origin and impact on revision work. *Instructional Science*, 39(3), 387–406.
- Liu, N. F., & Carless, D. (2006). Peer feedback: The learning element of peer assessment. *Teaching in Higher Education*, 11(3), 279–290.
- Lundstrom, K., & Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's own writing. *Journal of Second Language Writing*, 18(1), 30–43.
- Martins, D. (2008). Scoring rubrics and the material conditions of our relations with students. *Teaching English in the Two Year College*, 36(2), 123–137.
- Monahan, T., & Fisher, J. A. (2010). Benefits of 'observer effects': Lessons from the field. *Qualitative Research*, 10(3), 357–376.
- Pace University. (2018). Student body profile. Retrieved from <https://www.pace.edu/admissions-aid/undergraduate-admission/freshman/pace-student-body-profile>
- Pace University. (2019). ENG 201 Writing in the disciplines [Graduate catalogue course description]. Retrieved from <https://pace.smartcatalogiq.com/en/2017-2018/Continuing-Professional-Education/Course-Descriptions/ENG-English>
- Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, 35(4), 435–448.
- Rothman, J. (2012). The New Yorker's profiles. *The New Yorker*. Retrieved from <https://www.newyorker.com/books/double-take/the-new-yorkers-profiles>
- Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy & Practice*, 5(1), 77–84.

- Smith, H., Cooper, A., & Lancaster, L. (2002). Improving the quality of undergraduate peer assessment: A case for student and staff development. *Innovations in Education and Teaching International*, 39(1), 71–81.
- Swales, J. (2017). The concept of discourse community. In E. Wardle & D. Downs (Eds.), *Writing about writing: A college reader* (pp. 215–229). Boston, MA: Bedford/St. Martin's.
- Wardle, E., & Downs, D. (2017). *Writing about writing: A college reader*. Boston, MA: Bedford/St. Martin's.

Appendix A

Initial Survey

1. In my courses at this university, course requirements and grading systems tend to be clear.
Strongly Agree Agree Neutral Disagree Strongly Disagree
2. In my courses, exams and written assignments tend to be fair and reflect course content.
Strongly Agree Agree Neutral Disagree Strongly Disagree
3. In my courses, readings are effective in helping me understand how to approach written assignments.
Strongly Agree Agree Neutral Disagree Strongly Disagree
4. In my experience, professors offer rubrics (or well-articulated assignments) to define the expectations of written assignments.
Strongly Agree Agree Neutral Disagree Strongly Disagree
5. In my experience, professors use rubrics (or well-articulated expectations) to grade written assignments.
Strongly Agree Agree Neutral Disagree Strongly Disagree
6. Working and “talking it through” with fellow students has been essential to understanding a professor’s expectations on written assignments.
Strongly Agree Agree Neutral Disagree Strongly Disagree
7. Professor-directed peer feedback on written assignments has been valuable in my past courses.
Strongly Agree Agree Neutral Disagree Strongly Disagree
8. Peer feedback on written assignments can be valuable with the right instruction and guidelines.
Strongly Agree Agree Neutral Disagree Strongly Disagree

9. It is possible for students to grade each other's written assignments fairly and competently.
Strongly Agree Agree Neutral Disagree Strongly Disagree
10. Letter grades are an essential tool of education.
Strongly Agree Agree Neutral Disagree Strongly Disagree
11. Letter grades are an essential tool to evaluate students for employment and higher education.
Strongly Agree Agree Neutral Disagree Strongly Disagree
12. If I get less than a B, I feel like I have failed to complete the course's demands.
Strongly Agree Agree Neutral Disagree Strongly Disagree
13. I expect that the most common grade assigned to students at my university is:
A B C D F

Appendix B

Pregrading Survey

1. I understand the basic elements of this genre.
Strongly Agree Agree Neutral Disagree Strongly Disagree

2. I understand the requirements of this writing assignment.
Strongly Agree Agree Neutral Disagree Strongly Disagree

3. The writing element that I am most nervous about is (pick one):
Incorporating dialogue or quotes
Observational and descriptive writing
Biographical or historical writing
Expressing my subject's place in the world
Structure, grammar, and style
Entertaining my readers
Not having enough material
Having too much material
N/A: I am not nervous about this assignment

4. In order of importance: I learned the most about how to write this assignment through (drag to reorder):
Readings within the genre
Prior knowledge of the genre
Discussion of our rubric
Personal inquiries/readings
Peer review
Professor's lectures
One-on-one with the professor

5. I understand the genre well enough now to grade fellow students' work competently.
Strongly Agree Agree Neutral Disagree Strongly Disagree

-
6. I am nervous about other students grading my work fairly and competently.
Strongly Agree Agree Neutral Disagree Strongly Disagree
7. With students doing the grading, I expect the grades to be:
Higher
Lower
Same/Similar
8. It is possible for students to grade each other's written assignments fairly and competently.
Strongly Agree Agree Neutral Disagree Strongly Disagree
9. As I grade my fellow students' writing, I expect that I will learn (check all that apply):
More about my own approach to this writing assignment
Styles and techniques that I can incorporate into future writing
How to better evaluate peers thoughtfully and effectively
A better understanding of how others evaluate my writing
How to read similar texts more critically
About the person and subject matter the writer writes about
It will reinforce things I already know about writing and reading
I do not expect to learn anything valuable at all
10. I see educational value in having students grade each other's work.
Strongly Agree Agree Neutral Disagree Strongly Disagree
11. I am proud of the work I have been doing on this piece of writing.
Strongly Agree Agree Neutral Disagree Strongly Disagree

Appendix C

Postgrading Survey

1. I was nervous about having my peers grade my writing.
Strongly Agree Agree Neutral Disagree Strongly Disagree
2. I understood the basic elements of this genre enough to grade it.
Strongly Agree Agree Neutral Disagree Strongly Disagree
3. The written peer responses and comments (during grading) were done thoughtfully and with an eye toward making my essay better.
Strongly Agree Agree Neutral Disagree Strongly Disagree
4. The written response by my professor (during grading) was done thoughtfully and with an eye toward making my essay better.
Strongly Agree Agree Neutral Disagree Strongly Disagree
5. My peers graded my work fairly and competently.
Strongly Agree Agree Neutral Disagree Strongly Disagree
6. Grading my peers' writing, I learned (check all that apply):
More about my own approach to this writing assignment
Styles and techniques that I can incorporate into future writing
How to better evaluate peers thoughtfully and effectively
A better understanding of how others evaluate my writing
How to read similar texts more critically
About the person and subject matter the writer writes about
It will reinforce things I already know about writing and reading
I do not expect to learn anything valuable at all
7. It is possible for students to grade each other's written assignments fairly and competently.
Strongly Agree Agree Neutral Disagree Strongly Disagree

-
8. The peer-grading process was a valuable experience.
Strongly Agree Agree Neutral Disagree Strongly Disagree
9. I would rather have the professor be the only one to grade my work.
Strongly Agree Agree Neutral Disagree Strongly Disagree
10. In order of importance: I learned the most about how to write the profile essay through (drag to reorder):
- Readings within the genre
 - Prior knowledge of the genre
 - Discussion of our rubric
 - Personal inquiries/readings
 - Peer review
 - Professor's lectures
 - One-on-one with the professor
 - Peer grading/response
 - Professor's grade/response
11. I am proud of the work I have been doing on this piece of writing.
Strongly Agree Agree Neutral Disagree Strongly Disagree
12. It is possible for students to grade each other's written assignments fairly and competently.
Strongly Agree Agree Neutral Disagree Strongly Disagree
13. Would you suggest I do this again for the profile essay?
- No
 - Yes
 - Yes, with some changes
 - Yes, and you should do it with more assignments
14. What do you think was the most interesting or successful aspect of peer grading? (Written answer.)
15. What was the most challenging or unsuccessful aspect of peer grading? (Written answer.)

Appendix D

Student-Developed Rubric (Section 1)

| | Print-Worthy (A+) to Exceptional (A-) 25–23 | Strong (B+) to Developing (B-) 22–20 | Flawed (C) to Missing (F) 19–10 |
|---------------------------------------|---|--|--|
| A Broader Perspective | The essay offers a broader perspective—an angle on the subject’s place within a discipline, field, or idea. A general audience reader could both understand the subject’s place within this broader perspective and learn something new about the discipline, field, or idea. | The broader perspective is there, but it is nonspecific, thereby not shedding much or any new light on the subject’s place in the field or on the field. | The essay is focused more on personal interest issues, with little or no broader context. It lacks a connection to discipline, field, or idea. |
| A Biography or History of the Subject | The essay includes a biography or history of the subject. The biography is specifically organized and developed around the subject’s place in the discipline, field, or idea. A history of an aspect of the field is also included or indicated within the essay. | The essay includes a biography or history of the subject, or the field, but it is not as focused as it should be. | There is little to no understanding of the subject or field’s history. |

| | | | |
|--|--|--|---|
| | Print-Worthy (A+) to Exceptional (A-) 25–23 | Strong (B+) to Developing (B-) 22–20 | Flawed (C) to Missing (F) 19–10 |
| Quotes and Observations of the Subject | An interview took place, quotes were used directly, and the subject was described in action, in a compelling narrative way that allows the reader to see the person as they act in the world. | Quotes were used but were poorly integrated. Descriptions of the subject were missing or off topic. | One or more of the elements was missing and/or inadequate. |
| Organization, Clarity, and Grammar | Written without major grammatical errors, with language that fits the discourse community this essay might be written toward. The story is organized in a way that is not only understandable and clear but also compelling in the way the essay is formed and ends. | Few major grammatical errors. Organized in a clear and understandable way, but it could lack a compelling structure. | Error-riddled, or the diction does not fit the genre to an adequate degree. Or organization is confusing. |

| Category | Grade | Comments/Specific Feedback: |
|------------------------|-------|-----------------------------|
| Broader Perspective: | _____ | |
| Biography or History: | _____ | |
| Quotes & Observations: | _____ | |
| Organization, Clarity: | _____ | |
| Grade: | _____ | |

Copyrights

© JRW & Authors.

Copyright for this article is retained by the author(s), with first publication rights granted to the Journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (CC BY-NC-ND) (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Particelli, B. (2020). Student-led assessment: A small study on classroom rubric development and peer grading practices. *Journal of Response to Writing*, 6(1), 42–75.