The Effect of Peer Review on Student Learning Outcomes in a Research Methods Course

Jessica A. Crowe¹, Tony Silva¹, and Ryan Ceresola¹

Abstract
In this study, we test the effect of in-class student peer review on student learning outcomes using a quasiexperimental design. We provide an assessment of peer review in a quantitative research methods course, which is a traditionally difficult and technical course. Data were collected from 170 students enrolled in four sections of a quantitative research methods course, two in-class peer review sections, and two sections that did not incorporate in-class peer review over two semesters. For the two sections with peer review, content scheduled for the days during which peer review was used in class was delivered through the online course management system. We find that in-class peer review did not improve final grades or final performance on student learning outcomes, nor did it affect performance differences between drafts and final assignments that measured student learning objectives. Further, it took time away from in-class delivery of course content in course sections that used in-class peer review. If peer review is utilized, we recommend it be assigned as an out-of-class assignment so it does not interfere with in-class teaching.

Keywords
research methods, peer review, student learning outcomes

Student writing exercises are fundamental ways to achieve important institutional and course objectives (Grauerholz 1999; Hartmann 1992). Writing assignments are an active form of learning in which students can critically evaluate and reflect on course content. However, as noted by Althauser and Darnall (2001), the way many instructors structure writing assignments does not reflect the potential for writing to be an active learning experience. One way instructors can organize writing assignments to more effectively promote learning is to “scaffold” assignments into smaller pieces that can be reviewed by student peers. Scaffolding of writing assignments and peer review can increase students’ critical thinking and reading skills by starting with challenging tasks that require assistance from other students and the professor and then assigning additional challenging tasks. Massengill (2011) provides qualitative evidence that scaffolding a written assignment into several parts can aid in the intellectual leap from recall to application and analysis. Because assisted performance is achieved through scaffolding of projects and assistance from other students, peer

¹Southern Illinois University Carbondale, Carbondale, IL, USA

Corresponding Author:
Jessica A. Crowe, Department of Sociology, Southern Illinois University, Faner Hall, MC 4524, Carbondale, IL, 62901, USA.
Email: Jessica.Crowe@siu.edu
review can be used to maximize this pedagogical technique.

By using a quasieperimental design that compares student performance on course learning outcomes between sections of a quantitative research methods course that did not use in-class student peer review and sections of the same course that did incorporate in-class student peer review, we test the effect of in-class student peer review on student learning outcomes. There is little research about the effects of peer review in highly technical courses, such as methods courses. Our goal with the current research is to help fill that gap, assessing if in-class peer review helps students perform better when the student learning objectives (SLOs) are technical in nature. Technical courses, such as research methods, often involve memorization of procedures and appropriate standards for research. As such, although critical thought is still imperative, there are many topics and assignments on which there are clearly correct and incorrect responses (e.g., selecting topics appropriate for survey research and interpreting quantitative data). This is in contrast to many types of courses in which topics and assignments are more flexible, such as those exploring social issues. There may be differences between peer review outcomes in technical and nontechnical courses, since student knowledge (or lack thereof) regarding technical standards may negatively affect outcomes in technical courses, whereas unique insight—absent of technical knowledge—in nontechnical courses may prove beneficial rather than detrimental.

We attempt to answer the following question: Will incorporating in-class student peer review on scaffolded writing assignments lead to higher student achievement on SLOs and course grades and/or affect performance differences between drafts and final assignments measuring SLOs? To answer this question, we analyze SLOs that were assessed through the writing assignments. By doing so, we add to the research on active learning techniques, including writing across the curriculum, and uniquely evaluate the effect of in-class peer review on student performance independent of student characteristics, in a new setting.

STUDENT PEER REVIEW IN THE CLASSROOM

Student peer review, also referred to as peer editing, is a process in which students constructively evaluate the work of other students. At the college level, peer review can be used in a variety of courses and disciplines, ranging from English classes (Topping 2003) to business courses (Rieber 2006) to engineering courses (Carlson, Berry, and Voltmer 2005). In this process, students submit their work (e.g., written papers, homework) to peers, who review it and give feedback but who do not assign grades. Peer review in small groups of three to five students involves reading the work of several others. By assigning students to heterogeneous groups, students will be exposed to different writing styles. In addition, peer review can provide a network of peers, which is an important factor in explaining adult student retention (Thomas 2001; Tinto 2012).

According to Bruffee (1992:23–33), the ability to reason in a new discourse is an “artifact created by social interaction.” Therefore, it is one’s conversations with others that allows for the ability to think critically. One’s ability to write emerges from the ability to converse with others (Bruffee 1992). Peer review done in small groups can allow for such conversations to occur. Evidence suggests that, when used effectively, peer review can foster higher-order levels of learning, such as those represented by the upper levels of Bloom’s taxonomy; encourage students to develop professional behaviors that require the ability to reconcile multiple perspectives; and encourage students to develop the social skills needed to work in teams, including the ability to provide meaningful feedback and to accept peer critiques (Althauser and Darnall 2001; Carlson et al. 2005).

Several instructors who have incorporated peer review into their classes laud many benefits. Rieber (2006) compared assignments that were peer reviewed to student assignments that were not peer reviewed in a business communications class. He found that students did significantly better on assignments that were reviewed by other students than on assignments that did not undergo the peer review process. Researchers have evaluated the use of peer review by students with learning disabilities (MacArthur, Schwartz, and Graham 1991), low-achieving students (Karegianes, Pascarella, and Pflaum 1980), and high-achieving students (Topping 2003). Those studies and many others cited by Topping (2003) find that peer review can help students with a wide range of abilities to improve their written work.

Four reasons present themselves as possibilities for why peer review results in improved student outcomes. First, to participate in a peer review, students must complete their work at least one class before the assignment is due. This deters students from putting off writing until the last minute with
no time for revision. Second, because students review other students’ assignments to see if they followed directions and met the criteria for the assignment, the peer review process allows students a second chance to make sure they understand the assignment guidelines and follow them in their own papers. Students can learn just as much, if not more, from reviewing another student’s paper as they can from having their own work reviewed. Third, students may submit more polished drafts for peer review since they do not want to appear unintelligent to their peers. Fourth, peer reviewers tend to comment on what they do not understand or on how the guidelines are not being met. This can lead to discussions between the student being reviewed and the reviewer about what is needed or not needed in the assignment and can lead both students to reconsider the assignment (Rieber 2006).

Peer review may also offer substantial benefits for courses that focus on research methods and are thus more technical. Active learning assignments that allow students to offer feedback and discuss classwork might lead to increased student engagement and learning in methods courses, which are traditionally seen by students as passive and lecture based (Pfeffer and Rogalin 2012). Furthermore, in-class peer review sessions allow students to work on their projects during class time, and the structure of multiple drafts needing to be submitted for review ensures students stay focused on their work, both of which have been shown to be desirable for students in research-based courses (McKinney and Day 2012). Finally, small-group work that is focused on students delivering a particular final project can help students receive valuable feedback on their projects, which they can then incorporate into a more well-rounded final product (Albright et al. 2012).

Despite its potential as a rich pedagogical technique, there are some difficulties associated with student peer review. On the students’ end, some in the class may not understand the group’s purpose and what is expected of each student. Some students may be unable to read the work of others in an analytical manner that produces constructive feedback, while other students may fail to work collaboratively with others in their group (Spear 1988). Furthermore, redefining student roles with peer review can be difficult to grasp for some students. Some students may consider methods classes, in contrast to courses on topics like race or gender, to require expert instruction at all times and not the feedback of peers (Albright et al. 2012). On the instructor’s end, group activities may not be monitored properly. To overcome such pitfalls, instructors can scaffold large assignments into shorter assignments with very specific guidelines that build on each other. Additionally, instructors can direct peer review by giving reviewers a checklist that covers all aspects of the assignment guidelines. Applying review criteria to others’ writing can prompt its application to one’s own writing. Last, the instructor can also carefully monitor in-class peer review sessions.

HYPOTHESES

In this study, we test whether in-class student peer review has an effect on student performance independent of student characteristics. Drawing from the arguments presented that suggest student peer review can improve learning outcomes, we hypothesize the following:

Hypothesis 1: Students in the peer review sections will have significantly higher final grades than students in the sections without peer review.

Hypothesis 2: Students in the peer review sections will score significantly higher on student learning outcomes that are measured by their performance on assignments that undergo peer review than students in the non-peer review sections.

Hypothesis 3: Students in the peer review sections will have significantly higher final grades on SLOs measured by drafts and final assignments that undergo peer review than students in the sections without peer review.

We include the first hypothesis because we acknowledge that there may be a “spillover effect” from the peer review assignments to other course assignments and tests. In the process of providing peer review to others, students may revisit their notes or textbook to share more informed feedback; in this way, they may strengthen their understanding of course concepts. A similar process may occur when students read or hear peer review feedback from others; they may return to course materials to better understand what the peer review feedback means and to determine why their earlier draft needs improvement. Additionally, conversations between students during the process of peer review may strengthen students’ understandings of course concepts more generally, and this may influence student performance on non-peer review assignments and tests. We include the third hypothesis because we intend to capture the change in
performance that may be caused by the peer review process. As a result of the numerous possible benefits peer review could provide, we consider it important to examine final course grades, SLOs measured by written assignments, and the change in performance between drafts and final assignments measuring SLOs.

METHODS

Student Body and Course Design

To test the effect of student peer review of scaffolded writing assignments on student learning outcomes, we conducted a quasi-experiment within a quantitative research methods course offered at a small urban university in Texas. The university is a commuter-only college that serves residents of the city and surrounding counties. In the fall of 2012, there were 2,100 students enrolled in the university. Of those students, 40 percent were African American, 36 percent were Latino, 20 percent were white, and 4 percent classified themselves as another race or ethnicity. Women accounted for 61 percent of the student body. The median age was 27, with 16 percent of students over the age of 40.

Because students could pick which section of the course to take, rather than being randomly assigned to a section, the study qualifies as a quasi-experiment rather than a classical, randomized experiment. Quantitative Research Methods is a required course for both the sociology major and the applied arts and sciences major and requires a prerequisite of an introductory course in social science (e.g., Introduction to Sociology, Introduction to Psychology). Because it is a required course for the two majors, the course is mostly filled with students majoring in one of the two majors and is generally filled to capacity. A small number of students outside the two majors take it as an elective. One hundred and seventy-two students enrolled in four sections of the course during the 2012–2013 academic year. Of the students, 67 percent of students were female, 53 percent were African American, 25 percent were Latino, 20 percent were white, and the average age was roughly 26. The students in the four sections were representative of the larger student body, with a slightly higher percentage of women and African American students.

Quantitative Research Methods is designed to give the student a general understanding of sociological research methods. It examines many topics, including conducting surveys and other quantitative research, interpreting data, evaluating published research, and considering the ethics of research. By the end of the course, each student is expected to meet the course’s SLOs, including the two measured in the current paper: (1) the ability to formulate empirical research questions and (2) the ability to identify materials, research, and data relevant to research questions.

Student success in achieving course outcomes was assessed through a combination of multiple-choice, short-answer, and essay questions on three exams and a final research proposal that incorporated scaffolding. The final grade included the following: three exams, 10 percent each; the final research proposal, 25 percent; a presentation, 5 percent; five scaffolded written assignments, 4 percent each; questions over the weekly reading, 10 percent; and five in-class assignments (for the non–peer review classes) or five in-class peer review sessions, worth 2 percent each. The class assignments for the non–peer review sections were simple exercises, while those for the peer review sections were credit for participating in peer review. The research proposal was a final paper that applied what a student learned in the course to an original proposal for research. In the research proposal, the student had to come up with an original research question, two to four hypotheses related to the research question, and a detailed methods section outlining how the student would collect and analyze data to ensure valid results. Because the final research proposal was a daunting paper that most students were not accustomed to writing, the paper was scaffolded into five written assignments that built on the previous assignment.

The first and second written assignments, along with corresponding sections of the research proposal, directly measured how well a student met the two SLOs examined in the current paper. The first written assignment required students to create a research question and to write an annotated bibliography of 10 scholarly articles related to the research question. The second written assignment required students to formulate two to four hypotheses based on the research question from assignment 1 (or a revised research question) and to begin writing the outline of a methods section, including sample size, unit of analysis, operationalization of independent and dependent variables, and methodology. Students in all sections were required to complete all five written assignments and received extensive written feedback from the instructor on all assignments. Students in the peer review sections attended in-class peer review sessions for the first four written assignments, where, in groups,
they read over one another’s work and gave detailed feedback.

Quasi-experimental Design

Data used in this study come from four sections of the course taught by the same instructor over two consecutive semesters (fall and spring) during the 2012–2013 school year. The course used in-class student peer review over written assignments once in the fall term and once in the spring term. The other two sections (one in the fall and one in the spring) did not incorporate student peer review. For the two sections with peer review, content for the days with peer review were delivered through the course management system. Because less than 30 percent of the class was conducted online, it did not need to be advertised as a hybrid course. Students were made aware of the online component on the first day of class. No students dropped or switched sections after learning about the online component, so self-selection into sections based on the mode of course delivery did not occur.

For each semester, one section met twice a week during the day, and one section met once a week at night. To account for the fact that night students are more likely to work full-time than day students, the peer review class was taught during the day section in the fall and during the evening section in the spring. Likewise, the non–peer review class was taught during the evening section in the fall and during the day section in the spring. A coin was flipped to decide which section in the fall would use peer review and which one would not. Sections were capped at 50 students, and class sizes ranged from 34 to 50. For the two peer review sections, there were a total of 90 students, whereas the two non–peer review sections had 80 students.

The lead author taught all sections with the same course objectives, covered the same course material, used the same readings, completed the same major assignments and exams, and held students to the same standards with respect to comprehension of course material. The lead author posted assignments, lectures, rubrics, study guides, and sample proposals to the course management system for all sections. Students in all sections submitted assignments, aside from exams, through the course management system. Students in all four sections took exams in the classroom during the scheduled class time. Because the night sections were held for three hours once a week and the day sections for one hour and 30 minutes twice a week, students in all sections were allowed only one hour and 30 minutes to complete the first two exams, and two hours were allowed for the final exam. Therefore, students in all sections took their exams and submitted their writing assignments under identical conditions. In addition, exams and writing assignments submitted in all sections were evaluated with the same grading rubrics. The exact same exams and set of guidelines for writing assignments were used in all four sections. We did so because the university had only commuter students and had no history of test banks or a culture of students’ sharing test information with other students. To prevent cheating, the lead instructor created two versions of the exam with different question ordering on multiple-choice and true/false questions and also collected all exams after students looked at their grades and reviewed their answers.

For 75 percent of the course, both types of courses were exactly the same except for the timing and location of course lectures. However, two of the sections incorporated in-class student peer review sessions, while the other two did not. For the peer review sections, the lead author randomly assigned students to groups of three. Preventing self-selection helped ensure that students did not form groups with classmates they knew, which may have created a context in which students were uncomfortable giving informed feedback for fear of offending their friends. Randomly assigning students also increased the chances that the groups would be heterogeneous in terms of sociodemographic characteristics, personality, ability, and interest in the course, which may in turn lead to more meaningful feedback exchanges between group members. We chose to assign students to groups of three to ensure that all students received two sets of comments from their peers.

Students met four times throughout the semester in class to review and give feedback to other students’ written assignments. The written assignments were scaffolded pieces of the final research proposal. During class time, students broke up into their groups and read each other’s written homework. At the beginning of class, the instructor gave students detailed rubrics on what to look for in their peers’ homework and how to rate each item according to the rubric. For instance, for homework assignment 1, which required students to form a social science research question and to write an annotated bibliography for 10 scholarly sources related to the research question, students were provided with a list of questions to answer about their peers’ written assignment. All students in a group reviewed each other’s written assignments and
evaluated each work based on the questions provided. Questions for written assignment 1 included the following: Is the question a social research question? Is it something that can be empirically researched? Is it somewhat unique or do we already know the answer to the question? Are all 10 references scholarly? Is a complete citation provided for each reference? Does the summary of each reference provide the purpose of the study, the methods used, and the results? Students were required to discuss each question and then, as a group, rate each question on a five-point scale ranging from excellent to not present. They were encouraged to consult their text and notes for definitions of empirical, social research question, and scholarly.

After they were finished reviewing one student’s work, they continued with the next student in the group. While students were evaluating each other’s assignments, the instructor moved from group to group answering questions about the review process, the assignment, and course concepts, while also providing additional feedback on individual assignments. A teacher’s assistant took attendance and walked from group to group to ensure that students were on task. Peer sessions lasted for one hour and 30 minutes. This was equal to one full class period during the fall semester and a half of a class period during the spring semester. Students did not attend class for the remaining one hour and 30 minutes in the spring semester. Peer review sessions occurred prior to each homework’s due date. Therefore, students had time to incorporate peer feedback before turning in completed assignments for a grade. The students did not turn in their peer review feedback, so the peer review process itself was ungraded.

For the days when there were student peer reviews, class lectures were delivered through the online course management system. Lectures for those five classes were typed, and voice-recorded PowerPoint slides were created. The instructor provided the peer review sections both typed lecture notes and voice-recorded PowerPoint slides for the days of peer review and gave the sections that had regular lecture and no peer review regular PowerPoint slides and no lecture notes. To compensate for in-class discussion that normally occurs during class lecture, students were required to post questions over the week’s content on the discussion board housed on the course management system during weeks they had peer review. Students were required to post at least two questions for each week and answer a minimum of one other question for each week. The instructor commented on postings that were unanswered or needed clarification. By delivering lectures and discussions online, plenty of time was available in class for the student peer review sessions.

The lead author graded each written assignment—as well as the final research proposal—with a detailed rubric that explained how many points each assignment (or section) was worth and how points would be granted. This ensured grading was completed as objectively as possible. The rubric for homework assignment 1, for example, stated requirements that the research question was unique, concerned a social science topic, and could be researched quantitatively. The lead author performed all grading tasks, and grading was not double-blind; the instructor knew whose assignment she was grading (and by extension, if the student was in a peer review or non–peer review section), and students knew the instructor graded all assignments.

**Data and General Procedure**

The university’s institutional review board approved this study. The authors obtained the data from the instructor’s record of student grades, student exams, and grading rubrics of two of the written assignments as well as a student survey attached to the consent form. At the end of each semester, during the final exam, the instructor informed each class that she was conducting a research study. The instructor explained that although all exercises, assignments, and exams were mandatory for class credit, consent was needed for an individual’s course work to be included as research data. Students were instructed to sign the consent form, complete the attached survey, and turn it in to the sealed box when they turned in their final exam. Students were advised to turn a blank form into the same box if they did not want to participate. Students were told that the box would not be opened until after final grades were posted and that a student’s decision to participate or to withdraw from the study would not have an effect on one’s standing in the course or one’s final course grade. All but two students chose to participate in the study for a response rate of 98.8 percent, resulting in a total sample size of 170 participants (90 in peer review sections, 80 in non–peer review sections).

**Dependent Variables**

To answer our research questions, we assessed students’ final grades, two SLOs, and the change in
performance between drafts and final assignments measuring SLOs. Each student had a score for each SLO on a standard 0-to-100 percentile scale. A score was calculated by taking the student’s final score on all items used to assess an SLO, dividing by the total points allowed, and multiplying by 100. Change in performance between the first draft and the corresponding section in the final research proposal was assessed by subtracting the former from the latter.

**Final grade.** Each student’s final grade (on a standard 0-to-100 percentile scale) was calculated by adding all points earned on all exams and assignments in the class. Because the final grade included the scores of writing exercises, which in turn measured SLOs 1 and 2, it is being included as a dependent variable.

**Research question (draft).** The first written assignment required students to formulate an empirical research question. This was a first attempt at measuring SLO 1. The peer review sections incorporated in-class student peer review over the first written assignment, while the non–peer review sections did not incorporate student peer review.

**Research question (change in performance).** Change in performance of SLO 1, which required students to formulate empirical research questions, was evaluated by subtracting each student’s score of his or her research question draft from the score each received in the corresponding section of the final research proposal.

**Methods (draft).** The second written assignment required students to identify materials, research, and data relevant to their research questions. This was the first attempt at measuring SLO 2. The peer review sections incorporated in-class student peer review over the second written assignment, while the non–peer review sections did not incorporate student peer review.

**Methods (change in performance).** Change in performance of SLO 2, which required student to identify materials, research, and data relevant to their research questions, was evaluated by subtracting each student’s score of his or her methods draft from the score each received in the corresponding section of the final research proposal.

**Ethics.** This SLO was measured by a combination of multiple-choice and essay questions on the first exam. For the non–peer review sections, the instructor lectured in class on ethical standards when conducting social research. This same lecture was uploaded to the course management system for students in the peer review sections to view at their leisure.

**Independent Variables**

**Course type.** The major independent variable for the study was the use of student peer evaluation (hence referred to as peer review sections). Peer review sections were coded as 1, and non–peer review sections were coded as 0.

**Sociology major.** Because two majors required the course for graduation, a dichotomous variable was used to code for sociology majors (coded as 1) and non–sociology majors (coded as 0). This was taken from the student’s profile when he or she registered for classes.

**Year in school.** A student’s year in college was based on a five-point scale. Freshmen were coded as 1, sophomores as 2, juniors as 3, seniors as 4, and postbaccalaureate students as 5. Data were taken from the end-of-semester student survey.

**Student status.** Because many students attended the university on a part-time basis, a dichotomous variable was used to code for full-time students (coded as 1) and part-time students (coded as 0). Data were taken from the end-of-semester student survey.

**Sex.** A dichotomous variable was used to code for a student’s sex. Male was coded as 1 and female as 0. Data were taken from the end-of-semester student survey.

**Age.** An ordinal variable was used to code for a student’s age. Students who were 18 to 21 were coded as 1, 22 to 29 as 2, 30 to 39 as 3, 40 to 49 as 4, 50 to 59 as 5, and 60 and older as 6. Data were taken from the end-of-semester student survey.

**African American.** A dichotomous variable was used to code for whether or not a student identified as African American. Those who identified as African American were coded as 1; all others were coded as 0. Data were taken from the end-of-semester student survey.

**Latino.** A dichotomous variable was used to code for whether or not a student identified as Latino. Those who identified as Latino were coded as 1; all others were coded as 0.
were coded as 0. Data were taken from the end-of-semester student survey.

**Analysis**

We first analyzed data by examining descriptive statistics for variables and conducting independent-sample *t* tests for dependent variables. To control for other variables other than peer review that may “explain away” any initial differences between the peer review and non–peer review sections, we analyzed the effect of the major independent variable in combination with other independent variables on student grades and SLOs using ordinary least squares regression.

**RESULTS**

Table 1 presents the univariate and bivariate statistics for the independent and dependent variables based on course type. Chi-square tests show that students that attended the peer review sections were similar to those in the non–peer review sections with respect to sex, major, age, and year in school. However, the peer review sections had significantly more African American students and full-time students than the non–peer review sections (χ² = 4.419, *p* < .05; χ² = 4.620, *p* < .05, respectively). Comparison of means of the dependent variables show that students in the non–peer review sections significantly outperformed the students in the peer review sections with respect to final grades in the course (*t* = 1.75, *p* < .10). However, students in the peer review sections significantly outperformed the students in the non–peer review sections for the first written assignment, which was the first assessment of SLO 1 (*t* = 1.84, *p* < .10).

Table 2 contains the results of the relationship between course type on final grades and student learning outcomes while controlling for other characteristics of students. The regression findings do not find a significant relationship between course type and final grades, thus not supporting hypothesis 1, which stated that students in the peer review sections would overall perform better than the students in sections without peer review. A significant difference exists between the two course types for the draft of the research question (SLO 1), which was assessed by the first written assignment. Students in the peer review sections were better able to formulate empirical research questions than students in the non–peer review sections.

### Table 1. Univariate and Bivariate Statistics for Variables by Type of Course.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Non–peer review</th>
<th>Peer review</th>
<th>Chi-square and <em>t</em> test statistics*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 80)</td>
<td>(n = 90)</td>
<td></td>
</tr>
<tr>
<td>Independent variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sociology major (% yes)</td>
<td>45.00 —</td>
<td>33.00 —</td>
<td>2.43</td>
</tr>
<tr>
<td>Year in school</td>
<td>3.33 1.03</td>
<td>3.14 1.16</td>
<td>1.14</td>
</tr>
<tr>
<td>Full-time students (%)</td>
<td>57.00 —</td>
<td>74.00 —</td>
<td>4.62*</td>
</tr>
<tr>
<td>Sex (% male)</td>
<td>27.00 —</td>
<td>38.00 —</td>
<td>2.26</td>
</tr>
<tr>
<td>Ageb</td>
<td>2.55 1.23</td>
<td>2.75 1.32</td>
<td>0.79</td>
</tr>
<tr>
<td>African American (% yes)</td>
<td>45.00 —</td>
<td>61.00 —</td>
<td>4.42*</td>
</tr>
<tr>
<td>Latino (% yes)</td>
<td>30.00 —</td>
<td>20.00 —</td>
<td>2.28</td>
</tr>
<tr>
<td>Dependent variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final grade</td>
<td>77.08 14.63</td>
<td>72.69 17.65</td>
<td>1.75†</td>
</tr>
<tr>
<td>Research question (draft)</td>
<td>77.92 27.27</td>
<td>85.14 23.79</td>
<td>1.84†</td>
</tr>
<tr>
<td>Methods (draft)</td>
<td>76.74 20.74</td>
<td>71.67 26.44</td>
<td>1.38</td>
</tr>
<tr>
<td>Research question (change in performance)</td>
<td>10.79 34.61</td>
<td>4.10 27.00</td>
<td>1.41</td>
</tr>
<tr>
<td>Methods (change in performance)</td>
<td>-10.20 28.47</td>
<td>-11.64 23.31</td>
<td>0.36</td>
</tr>
<tr>
<td>Ethics</td>
<td>87.79 20.84</td>
<td>82.87 23.11</td>
<td>1.45</td>
</tr>
</tbody>
</table>

*a* A chi-square test was used to determine the significance of the difference in frequencies across course type for categorical data. An independent-samples *t* test was used for ordinal and continuous data.  
*b* Age is on a six-item scale (see Methods for details).  
† *p* < .10.  †*p* < .05.
Thus support exists for hypothesis 2, which states that students in the peer review sections will score significantly higher on student learning outcomes that are measured by their performance on assignments that underwent peer review than students in the non–peer review sections.

While an independent-samples $t$ test showed a significant difference for final grades between the two types of courses, this relationship is no longer significant when controlling for other individual characteristics. When performing regression, a student’s year in college, age, and race are significantly related to one’s final grade in the course. Senior-level students and older students scored significantly higher grades than junior-level and younger students ($B = 4.756, p < .001; B = 2.363, p < .05$, respectively). African American students scored significantly lower grades than students who were not African American ($B = -6.292, p < .05$). Because students could choose which section of the course to take, by chance, more African Americans enrolled in the peer review sections (61 percent of students identified as African American) than the non–peer review sections (45 percent of students identified as African American). For student learning outcomes, year in college is significantly related. Senior-level students scored significantly higher on drafts of the research question and methods than junior-level students. The relationship between year in college remains significant when measuring the change in performance from the draft and final proposal. Senior-level students had significantly higher changes in performance than junior-level students. However, course type was not significantly related to change in performance for student learning outcomes when controlling for students’ initial draft scores. Johnson (2005) suggests to also test two-wave panels by using the final score received as the dependent variable and the baseline score (i.e., the draft score in our model) as an independent variable. Testing the models in this way yielded no differences in which independent variables were significant. Thus we did not find

### Table 2. Ordinary Least Squares Regression Models Predicting Final Grades and Student Learning Objectives ($N = 170$).

<table>
<thead>
<tr>
<th>Independent and control variables</th>
<th>Final grade</th>
<th>Research question (draft)</th>
<th>Methods (draft)</th>
<th>Research question (change in performance)</th>
<th>Methods (change in performance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research question (draft)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>$-0.777^{***}$</td>
<td>0.076</td>
</tr>
<tr>
<td>Methods (draft)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Course type (peer review)</td>
<td>$-1.224$</td>
<td>$11.063^{**}$</td>
<td>$-1.245$</td>
<td>$1.459$</td>
<td>$-2.346$</td>
</tr>
<tr>
<td>(2.182)</td>
<td>(3.904)</td>
<td>(3.684)</td>
<td>(3.811)</td>
<td>(3.674)</td>
<td></td>
</tr>
<tr>
<td>Sociology major (yes)</td>
<td>1.580</td>
<td>-1.428</td>
<td>5.271</td>
<td>1.523</td>
<td>-1.040</td>
</tr>
<tr>
<td>(2.264)</td>
<td>(4.051)</td>
<td>(3.823)</td>
<td>(3.858)</td>
<td>(3.834)</td>
<td></td>
</tr>
<tr>
<td>Year in school</td>
<td>4.756^{***}</td>
<td>5.044*</td>
<td>6.017^{***}</td>
<td>8.467^{***}</td>
<td>5.130^{**}</td>
</tr>
<tr>
<td>(1.141)</td>
<td>(2.041)</td>
<td>(1.926)</td>
<td>(1.980)</td>
<td>(1.979)</td>
<td></td>
</tr>
<tr>
<td>Student status (full)</td>
<td>1.470</td>
<td>2.486</td>
<td>1.960</td>
<td>4.795</td>
<td>1.317</td>
</tr>
<tr>
<td>(2.479)</td>
<td>(4.435)</td>
<td>(4.185)</td>
<td>(4.226)</td>
<td>(4.175)</td>
<td></td>
</tr>
<tr>
<td>Sex (male)</td>
<td>-2.359</td>
<td>-7.015</td>
<td>-0.410</td>
<td>2.202</td>
<td>-2.941</td>
</tr>
<tr>
<td>(2.258)</td>
<td>(4.040)</td>
<td>(3.813)</td>
<td>(3.883)</td>
<td>(3.801)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>2.363*</td>
<td>0.548</td>
<td>0.376</td>
<td>2.080</td>
<td>0.903</td>
</tr>
<tr>
<td>(0.991)</td>
<td>(1.773)</td>
<td>(1.673)</td>
<td>(1.688)</td>
<td>(1.668)</td>
<td></td>
</tr>
<tr>
<td>African American (yes)</td>
<td>-6.292*</td>
<td>-6.809</td>
<td>-7.583</td>
<td>-0.26</td>
<td>-2.691</td>
</tr>
<tr>
<td>(2.821)</td>
<td>(5.047)</td>
<td>(4.763)</td>
<td>(4.833)</td>
<td>(4.786)</td>
<td></td>
</tr>
<tr>
<td>Latino (yes)</td>
<td>-0.684</td>
<td>-3.655</td>
<td>-3.680</td>
<td>4.984</td>
<td>-0.368</td>
</tr>
<tr>
<td>(3.116)</td>
<td>(5.574)</td>
<td>(5.260)</td>
<td>(5.313)</td>
<td>(5.251)</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>.245</td>
<td>.127</td>
<td>.115</td>
<td>.451</td>
<td>.264</td>
</tr>
</tbody>
</table>

*Note: Unstandardized coefficients reported with standard errors in parentheses. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$ (using a two-tailed $t$ test).
support for hypothesis 3, which stated that students in the peer review sections would have higher changes in performance on SLOs measured by drafts and final assignments that underwent peer review than students in the sections without peer review.

DISCUSSION

In this study, we examined how in-class student peer review of written assignments affects student learning. By performing a quasi-experiment on four sections of a quantitative research methods course taught over two semesters, we were able to test how students in sections that incorporated in-class student peer review performed on student learning outcomes compared to students in sections that did not use peer review. Through the use of multivariate regression analysis, we were able to test the relationship between the two types of courses and student performance while controlling for a number of other relevant student characteristics.

Students in the peer review sections performed better on a student learning outcome (SLO 1) that was assessed by a draft writing assignment that underwent student peer review than students in the sections who completed the writing assignment without peer review. No significant difference in outcomes was observed with the second SLO. Importantly, when the same SLOs that were measured by the scaffolded writing assignments were assessed by the same parts in the final end-of-semester research proposal, no significant difference in performance existed between the two types of courses. Because the instructor used scaffolding of the final research proposal in all four sections, these findings suggest that students in the non–peer review sections incorporated instructor feedback and were able to perform as well as the students in the peer review sections by the end of the semester. Further, there was no significant difference between the peer review and non–peer review sections regarding the change in performance between the drafts and final assignments that measured SLOs and underwent peer review. Thus, although there was less room for improvement between the draft and final assignment measuring SLO 1 for students in the peer review sections because they performed higher on the draft assignment measuring SLO 1, overall, the peer review did not affect change in performance on assignments measuring SLOs.

Even though there were no significant differences between the two types of courses with final grades and change in performance for SLOs measured by written assignments, the fact that students in the peer review sections performed more strongly on an SLO measured by an earlier draft of a writing assignment than students in non–peer review sections suggests that peer review does have value. This may especially be the case for students who struggle to understand course concepts more than their peers. The lead author and instructor provided comments on all drafts, so it is possible that students dismissed peer review feedback as irrelevant and unnecessary; future research should explore whether peer review is effective in courses in which the instructor is not able to provide comments on all assignments.

The finding that students in the peer review sections performed more strongly on the draft assignment measuring SLO 1 than students in the non–peer review sections suggests peer review utility may depend on the type of assignment in addition to the type of course. The assignment measuring SLO 1 may have been more conducive to peer review than the assignment measuring SLO 2. Formulating an empirical research question (SLO 1), arguably, is a more flexible assignment that involves well-defined, but broad, standards of research. Identifying materials, research, and data relevant to research questions (SLO 2), however, is much more technical and also more specific to each student’s project. Thus, it may be the case that peer review has more utility on assignments that involve the application of broader standards as opposed to more technical and specific ones.

Notably, we did not observe statistically significant differences between the two types of courses with regard to assessment of material that was covered solely through content presented online for peer review sections and in class for the non–peer review sections. For example, material covering ethics was placed online to make time for in-class peer review. Even though the peer review sections performed lower on test items covering the ethics of research than the non–peer review sections that had covered ethics in class (see Table 1), this difference was not significant. Unfortunately, many students in the peer review sections did not view the online lectures (e.g., 14 views of one 15-minute lecture clip for a class of 40). Despite repeated reminders to view the videos, a study guide that stressed the importance of knowing the material presented in the videos, and greater flexibility on when to view the online lectures, more students attended classes (on the days of scheduled class) than viewed the online lectures. This may be an issue of time, as the peer review sections were required to prepare drafts of assignments for the in-class peer review plus view online lectures out of...
class. While these students may have focused more energy on their papers, it left them less time to devote to content not covered in class. The absence of statistically significant differences, however, suggests that it may not be detrimental to place content online for some types of material in certain technical courses. However, given that some students will not view online material, faculty must carefully consider whether or not to move content online to make time for in-class peer review. It is possible that assigning peer review as an out-of-class homework exercise may be more effective than in-class peer review, as it would not be necessary to transfer in-class teaching material online. In any case, our findings indicate the need for further research to examine the utility of presenting content online to make time for other in-class activities.

It is important to consider the context of the current study. This study was conducted on students who attend a small commuter public university in an urban setting. Most students at the university are transfer and nontraditional students. Because transfer and nontraditional students are more likely than traditional students to work, it is possible they did not have enough time to provide quality feedback on assignments. Further, because many students commuted, they may not have felt as much pressure to provide quality feedback, as they did not live among their peers, unlike many traditional noncommuter college students. Given this limited demographic of students, the arguments presented would be further enhanced by similar experiments performed in other classes taught at other types of institutions of higher education.

Additionally, future research should examine whether peer review is more effective for courses that explore more abstract concepts than methods courses, such as classes about social issues. Research methods courses are highly technical, so one possible reason the in-class peer review was not very effective is because more guidance was needed from the instructor to more fully understand technical concepts. If students did not fully understand course concepts, then their feedback would not have been very helpful. Relatedly, the quality of the peer review given by students in this study may not have been high enough to help others. Although peer review quality is an important consideration for peer review in all contexts, it may be of special concern in courses that are highly technical, such as methods courses. The results of this study may have been different had the peer review feedback been graded and monitored for accuracy. It is possible that ungraded and unmonitored peer feedback may actually hamper learning in technical courses if incorrect feedback is given. Future research should explore the impact of peer review in different types of courses, while also taking into account the quality of peer review feedback by utilizing grading criteria for this feedback.

Future research could modify other elements of this study as well. Although students may greatly benefit from face-to-face conversations during the peer review process, some students may feel uncomfortable giving critical (and potentially higher-quality) feedback with their classmates present. Other research should explore the effectiveness of peer review when (1) the peer review comments are made outside of class and away from other group members or (2) when the instructor ensures the peer review is anonymous. An instructor could ensure the latter scenario by collecting students’ homework assignments, making copies and removing the students’ names, randomly assigning the students’ assignments to classmates, and collecting the assignments—with peer review comments—and later distributing them to the writer of the assignment. In addition, future studies should test the effectiveness of peer review using double-blind grading conditions; this would prevent any unintentional bias from subtly influencing the grading process. Last, although we randomly assigned students to groups to increase the likelihood that groups would be heterogeneous along several dimensions, and thus more beneficial to students, future researchers may seek to assign students to groups nonrandomly to ensure groups are as heterogeneous as possible. The results of the current study may have been different had we maximized group heterogeneity by nonrandomly assigning students to groups.

Although peer review should be carefully structured to maximize its benefits, there should be a critical discussion about what constitutes a benefit. For many students, peer review may be one of the only ways to actively participate in the learning process and engage with fellow students about course material. In this way, peer review may strengthen social skills, enhance confidence, and reaffirm that students can offer important feedback, demonstrating that the teacher—the formal authority figure—is not the only person who can contribute to the intellectual development of others. These kinds of benefits cannot be measured through graded performance, and yet they may qualitatively enrich a student’s learning experience. This may be especially important for nontraditional students and students of color, who are
underserved and may not have had the opportunity to undergo these potentially rewarding experiences. Future studies should consider these kinds of difficult-to-measure benefits.

Based on the results of this experiment, we caution instructors considering using in-class peer review for courses that are technical in nature, such as methods courses, because it did not improve final grades or final performance on SLOs. Further, it took time away from in-class delivery of course content in course sections that used peer review. As a result, students’ performance suffered on some assessment of material that was covered solely through content presented online for their sections than the non-peer review sections that had all content covered in class. Future studies should examine whether peer review is effective (1) when assigned as an out-of-class exercise, where it does not take time away from in-class delivery of teaching material; (2) in different types of courses that are not as highly technical as methods courses; and (3) when the peer review itself is graded and monitored for accuracy and helpfulness.

**EDITOR’S NOTE**

Reviewers for this manuscript were, in alphabetical order, Christopher Stout, Jennifer Strangfeld, and Alison Wisecup.

**REFERENCES**


Carlson, Patricia, Frederick Berry, and David Volتمر. 2005. “Incorporating Student Peer-review into an Introduction to Engineering Design Course.” Presented at the ASEE/IEEE Frontiers in Education Conference, October 19–22, Indianapolis, IN.


**AUTHOR BIOGRAPHIES**

Jessica A. Crowe is an assistant professor of sociology at Southern Illinois University Carbondale. She has recently...
published in *Assessment and Evaluation in Higher Education* and *Patterns of Prejudice*.

**Tony Silva** is a second-year MA student at Southern Illinois University Carbondale. His primary research interests include sexualities and gender, and his secondary research interests include environmental and rural sociology. He is currently interviewing rural men who identify as straight and have sex with other men to understand how they perceive their sexual identity and sexual behavior.

**Ryan Ceresola** is a second-year PhD student at Southern Illinois University Carbondale. He is currently pursuing a mixed-methods dissertation project on political corruption in the United States. His work has recently appeared in *Assessment and Evaluation in Higher Education* and *Patterns of Prejudice*.