A peer evaluation training results in high-quality feedback, as measured over time in nutritional sciences graduate students

Olivia S. Anderson, Noura El Habbal, and Dave Bridges
Department of Nutritional Sciences, University of Michigan School of Public Health, Ann Arbor, Michigan

Submitted 19 August 2019; accepted in final form 16 March 2020

Anderson OS, El Habbal N, Bridges D. A peer evaluation training results in high-quality feedback, as measured over time in nutritional sciences graduate students. Adv Physiol Educ 44: 203–209, 2020; doi:10.1152/advan.00114.2019.—Peer evaluation skills are not typically taught to students, yet they are expected to provide high-quality feedback to their peers. Gameful learning, a pedagogy supporting student-driven learning, can further reinforce the development of peer evaluation skills, if students are motivated to improve upon them. To better understand the effects of a peer evaluation training on the quality of student-generated peer evaluations, we scored peer evaluations from two cohorts taking a graduate-level nutritional sciences class using gameful learning pedagogy. The intervention group completed a peer evaluation training before engaging in peer reviews, while the control group did not. The training included two readings, a video, and reflection questions. The peer evaluations submitted by both the intervention and control groups were assessed on a validated rubric. The peer evaluation training had a positive effect on the quality of the submitted peer evaluations. The intervention group had a 10.8% higher score on its first submitted peer evaluation compared with controls (P = 0.003). The intervention group improved the quality of its future submissions by a further 8.9%, whereas the controls did not continue to improve substantially (P < 0.001). Overall, peer review training enhanced the quality of peer evaluations and allowed students to develop professional skills that they can utilize in any biomedical profession. Our results highlight the importance of peer evaluation training in combination with repeated practice and student-driven learning brought forth by gameful learning pedagogy in improving the quality of evaluations and developing professional skills.

INTRODUCTION

Peer evaluation is an active learning opportunity for students to directly engage in the exchange of feedback on their course assignments (6). During participation of a peer evaluation process, a student practices transferable professional skills, including collaboration, active listening, communication, and critical assessment (15, 25). With the exchange of formal feedback, recipients of the evaluation have an opportunity to improve their work or performance, but also must learn to regulate the information they receive and assess whether it could increase or decrease the value of their work.

Students are often required to complete assignments that emphasize the use of professional attributes needed to conduct a peer evaluation, like developing quality feedback, communicating in a professional manner, or taking autonomy on providing feedback, for example (28). However, instructors tend to expect that these professional skills have already been developed by the students, or are innate, and thus do not spend time adequately training students on how to effectively develop and utilize these types of transferable skill sets (26, 33). Although substantial academic research has focused on providing peer reviews in writing classes where English is the second language, studies on peer evaluation opportunities are lacking in the basic sciences (11, 22). The ability to develop quality feedback is important to make peer evaluation exercises effective, but what is equally important is understanding how and what type of feedback to incorporate that will make the work stronger. The opportunity to improve one’s work may be lost if the students are not given the opportunity to improve their work product given the feedback, and, likewise, the ability to generate and disseminate quality feedback cannot be gained. As such, the peer evaluation process would benefit both the recipient and evaluator, if they are given multiple opportunities to engage in such practice (7, 10).

A pedagogical approach that allows for flexibility of assignment types and offerings in which students can engage is called gameful learning. In gameful learning, multiple assignments are designed by an instructor, allowing the students autonomy in learning by choosing and completing course work that supports their interests and career goals (1). As the instructor designs the curriculum, there is an opportunity for him/her to create diverse assignments that help facilitate the student’s development of course competencies, as well as practical skills like peer evaluation. As such, peer review training and engaging in multiple opportunities to submit peer reviews allows students to become better critics and to improve the quality of their final submitted work (22, 26, 27).

In this paper, we describe a peer evaluation assignment consisting of a peer evaluation training for students, followed by multiple opportunities for students to engage in peer evaluation assignments. The peer evaluation was done within the context of a graduate-level, foundational nutritional sciences course based on the gameful learning teaching approach. Our aims were to arm students with a framework of best practices for basic peer evaluation and to give students multiple opportunities to practice generating peer evaluation feedback to improve over the course of the semester. To our knowledge, this is the first time that peer evaluation training and its outcomes in regard to the quality of the peer feedback have been reported in a biomedical sciences course.
Methods

Course description. The course in which we implemented the peer evaluation training and assignment as course assessments is called Nutritional Sciences: Principles of Nutrition. This course is required for all graduate students enrolled in the Nutritional Sciences program at the University of Michigan, Ann Arbor, including Master of Public Health, Master of Science, and Doctoral students. Additionally, a handful of graduate students from other programs, such as kinesiology, typically take the course. This course is completed in the first semester of the graduate program and, currently, is the only course using a gameful learning pedagogy approach in the University of Michigan School of Public Health. Briefly, required assignments are still valuable with gameful learning to ensure students are able to master core competencies; thus as instructors we made sure certain students would not solely focus on optional assignments without demonstrating mastery of the courses’ core competencies. In addition to required assignments, students were encouraged to choose from a variety of optional assignments, such as peer evaluation, in-class presentations, written reports, and formulating review questions. The detailed list of assignments, their descriptions, and points were explained in the syllabus. More detail in regard to the point system and GradeCraft, a web-based learning management system for gameful pedagogy, is described elsewhere (5). The focus of the course is gastrointestinal physiology, endocrinology, biochemistry, and metabolism, all discussed in the context of how the body handles macronutrients. There were three instructors for this course. The primary instructor teaches nutrient biochemistry, metabolism, and endocrinology; the co-instructor taught gastrointestinal physiology; and a graduate student instructor assisted in grading and teaching. The three instructors are the investigators of this study.

Study participants. The intervention group consisted of the 2018 cohort. The intervention cohort engaged in a peer evaluation training and, subsequently, had the option to complete peer evaluation assignments on student presentations throughout the semester. The control group was the 2017 cohort who could participate in peer evaluation on student presentations but did not have a peer evaluation training available to them.

Demographics of the students were obtained from the University’s registrar. Demographic data that were collected included sex, undergraduate degree, current degree program, race/ethnicity, and TOEFL (test of English as a foreign language) score (if applicable).

Peer evaluation training and assignment. The peer evaluation was one of six optional assignments available for the students to choose from. This model allows for student autonomy in choosing their activities throughout the course, including peer reviews, the focus of this paper. A separate optional assignment in which the students could choose to engage was an in-class presentation. The in-class presentation was the assignment on which the students peer evaluated each other. The peer evaluation assignment was set up in a way that, if students wanted to participate in it as one of their assignments, they had to first complete a peer evaluation training. Even though students were not required to complete peer review training, 97% of students in the intervention group elected to complete the module.

The peer evaluation training focused on the review of oral presentations and had three components to it. First, the students were assigned two readings. The first reading covered best practices of student presentations to help the student understand aspects of an effective presentation (9). The second reading helped the student learn about strategies to effectively peer evaluate a work product (12). The second requirement for the peer evaluation training was for the students to watch a 9-min video of a student presentation (external to the University of Michigan) and to provide feedback given the peer evaluation assignment rubric, meant for student use, that they would use for the actual peer evaluations for the in-class presentations (APPENDIX A). Last, after the readings and practice peer evaluation, the students were asked to reflect on their process by answering the following three questions: 1) What were two takeaways that struck you as important as you did the readings? 2) What will you strive to implement as you engage in peer evaluations in this course? and 3) Professionally speaking, do you think peer evaluation skills are important? The co-instructors provided written feedback on the quality of each student’s peer review practice based on a rubric similar to the one used in our final evaluation. The feedback included both comments and a score. However, the rubric was not identical to the final rubric in APPENDIX B, which was modestly refined during our internal validation and scoring process.

After completing the training, the students had the option to participate in up to five peer evaluations over the course of the semester. Each peer review was on a distinct presentation.

The students utilized an assignment rubric (the same used in the training; APPENDIX A) as a template for generating their feedback and to provide as a document to their peer. The co-instructor graded the quality of all generated peer feedback using a rubric meant for grading purposes (instructor use only; APPENDIX A) for the entirety of the semester. Note that the control group did not have access to the peer evaluation training, but did receive instructor feedback on the quality of each peer evaluation completed. The rubric was adapted from another course taught by one of the study investigators (4) and specifically provided guidance to the instructor to grade the quality of the students’ feedback based on the following subcomponents: 1) relevance of the feedback in regard to the assignment requirements; 2) specificity of the feedback; 3) how realistic the feedback was; 4) general use of a positive tone; and 5) professional quality of the feedback.

Qualitative methods. The text answers that the students provided from the reflection questions during the peer evaluation training were analyzed for major themes. One study investigator reviewed a sample of 30 text answers by reviewing each individual sentence and going back and reviewing each paragraph to refine the themes that emerged from the sentences. The text answers were then reviewed similarly by a second investigator. The themes that were uncovered by each investigator were discussed together in person until agreement was reached on the major themes that emerged from the reflection questions.

A voluntary, follow-up survey was provided to the presenters asking, “After receiving the feedback from your peers, do you recall any initial reactions, i.e., were the evaluations helpful, did the evaluations negatively impact your work or your feelings/attitudes, was the feedback what you expected, etc.?”, to help understand their emotional reactions to the feedback. The answers were analyzed for major themes similarly to the methods for analyzing the reflection questions described in the previous paragraph.

Rubric validation and assessment of peer evaluation quality. Proceeding the completion of the 2018 course, the investigators utilized the grading rubric to blindly reassess the quality of peer evaluations generated by both the control and intervention groups (n = 65 and n = 100, respectively, from 51 and 59 students). Before assessing all peer evaluations, the investigators underwent a comprehensive process to validate the grading rubric.

The study investigators adapted the Delphi method to validate the rubric (2). First, all three investigators used the existing grading rubric to assess six randomly selected peer evaluations: three from the control group and three from the intervention group. The investigators met to compare and discuss the quantitative and qualitative feedback they recorded. Discrepancies in grading were noted. Minor updates were made to the rubric after the discussion. The investigators reassessed the same six peer evaluations and again met to compare and discuss the feedback they recorded, and interrater reliability was assessed in R with the irr package (version 0.84.1). The investigators made minor adjustments to the wording in the rubric to help
facilitate further reviews. A separate random sample of six peer evaluations was selected, again, three from the control group and three from the intervention group. The updated rubric was used again to assess the new random sample. The investigators met to compare and discuss the feedback recorded and made final adjustments to the wording of the rubric. Interrater reliability was calculated for each round of assessment of the randomly selected samples. In the initial review round, we calculated an intraclass correlation coefficient of 0.3, improving to 0.57 after discussion and minor modifications of our rubric. This represents moderate agreement between raters (16). After the interrater reliability was at a moderate level and the discrepancies across the grading were resolved, the three investigators were randomly assigned to each grade one-third of the sample. A post hoc analysis to test for coder bias determined no significant moderating effect of the grader ($P = 0.69$).

**Quantitative methods.** Scores were evaluated both as aggregated total scores and as the individual components. All analyses were completed using R statistical software (version 3.5.0; Ref. 27). For pairwise tests, nonparametric Mann–Whitney $U$ tests were used, as the intervention group data were not normally distributed ($P = 0.05$ from a Shapiro–Wilk test). For longitudinal analyses, mixed linear models were constructed with random intercepts for each student. Students completed various numbers of peer reviews over the semester (see Table 1). Since the peer evaluation training was an opt-in intervention, we analyzed the improvements in peer review scores as intention to treat, not excluding students who completed less than five reviews. These analyses were done using the lme4 package (version 1.1–21; Ref. 3). For all analyses, statistical significance was designated as $P < 0.05$.

The study protocol was approved by the Institutional Review Board at the University of Michigan as exempt (HUM0000145801). Student consent was obtained through authenticated e-mail (control group) and through an electronic consent form (intervention group).

**RESULTS**

**Demographics.** There were no significant differences in the proportion of sex, ethnicity, age, and number of peer evaluations that the students conducted between cohorts (Table 1). Of the students who had the peer evaluation training available to them, the majority of the intervention cohort engaged in the opportunity ($n = 54$ or 97%), even if they never participated ($n = 1$ did not do a peer evaluation of the 54 who participated in the training) in a peer review assignment later throughout the course. Of note, there were a total of 7 oral presentations in the intervention group and 22 oral presentations given in the control group, so everyone who did a peer evaluation(s) did not necessarily also do an oral presentation. The range of peer reviews received by a presenter was 19–43.

**Qualitative analysis.** After engaging in the peer evaluation training, students articulated several strategies that they would apply to their own written peer evaluations and how practicing peer evaluation would be beneficial to them as they enter a professional setting. Two major themes that occurred throughout the written reflection responses at the conclusion of the peer evaluation training included students acknowledging that 1) when giving feedback they will incorporate positive aspects of the peer’s work (i.e., what peers did well), along with the areas on which to improve ($n = 32$ of written responses; 59.3%); and 2) peer evaluation promotes self-reflection for both the evaluator and the recipient of the feedback ($n = 30$ of written responses; 55.6%). Other themes that emerged to a lesser extent included the following: 1) giving feedback in a timely manner; 2) preparing for peer evaluation by reading the assignment rubric; and 3) providing specific feedback rather than general feedback.

Three of the seven presenters (43% response rate) responded to a follow-up survey asking about their initial reactions to the peer feedback. After conducting a thematic analysis of the open-ended question, two main themes were agreed upon, including the following: 1) teaching about feedback as a skill led to a better experience, helping to receive feedback in a positive manner; and 2) the received feedback was overall constructive compared with previous experiences.

**Effects of training on initial evaluation quality.** The students who engaged in the peer evaluation training had, on average, higher scores on the quality of their written feedback on the first peer evaluation assignment compared with the cohort who did not have access to the training (Fig. 1A). For the first evaluation, we observed a 10.8% increase in the overall review score for the students who completed the training ($P = 0.003$ via a Mann–Whitney $U$ test). Of the five criteria in the rubric, positivity and professionalism scored significantly higher on the first review for the students who had the training versus those who did not ($P = 0.02$ for each). The trend toward increased scores was generally true across all subcomponents of our rubric (Fig. 1B). Even though there were no significant differences in age, sex, or ethnicity between the cohorts (see Table 1), we tested whether including demographic factors modified the strength of the effect. As shown in Table 2, the beneficial effect of training in the first review score was consistent after adjusting for sex, age, ethnicity, or all three covariates combined (Table 2).

**Effects of training on improvement in evaluation quality.** We next asked if the initial training and feedback throughout the semester had an effect on whether scores pertaining to the quality of the student feedback continued to improve. As shown in Fig. 1C, the cohort of students who did not receive training did not substantially improve in their evaluation quality after five subsequent attempts (0.8% increase), but even though students who took training had a 10.8% increase in

**Table 1. Participant breakdown**

| Cohort 1 | Cohort 2 | $P$ Value |
|----------|----------|-----------|
| $n$      | 51       | 59        |           |
| Sex      |          |           |           |
| Female   | 46 (90)  | 49 (83)   | 0.42      |
| Male     | 5 (10)   | 10 (17)   |           |
| Age, yr  |          |           |           |
| 20–25    | 35 (67)  | 34 (58)   | 0.49      |
| 25–30    | 13 (25)  | 20 (34)   |           |
| 30+      | 3 (6)    | 5 (8)     |           |
| Ethnicity|          |           |           |
| White    | 28 (54)  | 38 (64)   | 0.29      |
| Asian    | 4 (8)    | 10 (17)   |           |
| Hispanic | 3 (6)    | 4 (7)     |           |
| Black    | 2 (4)    | 1 (2)     |           |
| Peer evaluations, no. |  |  |  |
| 1        | 1 (2)    | 2 (3)     | 0.09      |
| 2        | 3 (6)    | 3 (5)     |           |
| 3        | 1 (2)    | 2 (3)     |           |
| 4        | 5 (10)   | 11 (19)   |           |
| 5        | 23 (45)  | 33 (56)   |           |
| Peer evaluation training | NA | 57 (97) | NA |

Values are $n$, no. of participants (with percentage of participants in parentheses). NA, not applicable. $P$ values are from chi-squared tests, excluding when parameter is not indicated.
their initial score, they had a further 8.9% increase after repeated attempts. Via pairwise analyses using mixed-linear models, the intervention group had significantly higher overall scores ($P < 0.001$) and significant improvements in their scores over the course of the semester relative to the control group ($P = 0.05$). The improvement over the semester was still observed after adjusting for demographic factors (Table 2).

Another way to consider these effects, in the preintervention year (control group), only 38% of students improved the quality of their peer evaluations by any amount, but in the intervention year that number nearly doubled to 67% ($P = 0.02$ by Fisher’s exact test). The improvements in peer evaluation scores indicate that the intervention prompted improvements throughout the semester, over that of repeated practice alone.

DISCUSSION

To our knowledge, this is the first report of a peer evaluation training implemented in a biomedicines course. Other disciplines, such as teacher education or English as a second language (ESL) that more often report training effectiveness focus on the value of the peer evaluation to the recipient. For example, the quality of revisions incorporated and the quality of the recipient’s revised work were previously measured as outcomes (17, 20, 21). In our case, our focus was on the quality of the peer feedback itself. The intervention group that participated in the peer evaluation training before engaging in the peer evaluation assignments had higher quality feedback provided to their peers than the control group in this study.

Peer evaluation trainings reported in the literature typically comprise a combination of, or all of, the following: a case scenario modeled off of what the students will eventually assess, a rubric, discussion with a facilitator, and a student reflection on what they are learned (19, 26). Such trainings on how to approach a peer evaluation provide students a chance to practice and become confident to independently produce quality feedback and, subsequently, give it to a peer. The scaffolded nature of our peer evaluation training and assignments—readings, practice peer evaluation, reflections, multiple peer evaluations with feedback—was strategic, so the students had allocated time and space to reflect on the process that they undertook during the training, and then to apply their learnings from the dedicated training to actual, graded peer evaluations over the course of the semester. Research shows that, after engaging in peer evaluation trainings, students exhibit a greater sense of what is expected of them, improvements in the quality of feedback produced, as shown here, the amount of feedback given, and positive perceptions reported by the recipient who gets the feedback from the peer, leading to more revisions incorporated into their work (22, 35, 36). We noted increased participation in peer reviews in our intervention group, suggesting that training may promote higher participation in op-

Table 2. Multivariate analyses of the effects of training on the first evaluation and improvements over the semester from the first to the last evaluation score

| Outcome                  | Model       | Effect of Training Estimate, mean (SD) | $P$ Value |
|--------------------------|-------------|---------------------------------------|-----------|
| First review score       | Training    | 0.90 (0.27)                           | 0.001     |
|                          | + Sex       | 0.89 (0.27)                           | 0.002     |
|                          | + Age group | 0.91 (0.27)                           | 0.001     |
|                          | + Ethnicity | 0.84 (0.27)                           | 0.003     |
|                          | + Sex, age group, ethnicity | 0.83 (0.28) | 0.004 |
| Improvement over the semester | Training | 0.59 (0.33)                           | 0.07      |
|                          | + Sex       | 0.64 (0.33)                           | 0.05      |
|                          | + Age group | 0.60 (0.33)                           | 0.07      |
|                          | + Ethnicity | 0.60 (0.35)                           | 0.09      |
|                          | + Sex, age group, ethnicity | 0.65 (0.35) | 0.07 |

Scores are out of 10 points.
Another crucial aspect of the peer evaluation assignment was that, for both the intervention and control group, this was an assignment the student could choose to engage in multiple times throughout the semester due to the nature of the gameful learning pedagogical approach. The ability to develop feedback more than once meant that students would get an evaluation back from the instructor about the quality of their feedback several times during the semester, enabling the students to receive feedback themselves and incorporate strategies to improve the caliber of their peer evaluation (38). This process supports the emergent theme from the written reflections that peer evaluation fosters self-reflection for both the peer being evaluated and for the person who is developing feedback that is meant to be effective and meaningful to that recipient. Furthermore, this learning opportunity supports the development of professional competencies for those who master skills at different rates (30).

Because this was an optional assignment in the context of the gameful learning teaching approach, the students could choose whether peer evaluation was a skill that they wanted to improve upon and found useful for their future career goals (14). The gameful learning approach gives students autonomy over the course work in which they engage, promoting intrinsic motivation toward their professional development (1, 31). Thus, given multiple opportunities to improve the quality of their feedback, students start to understand their abilities surrounding the professional skill set that is needed for effective peer evaluation (18). The student autonomy in conjunction with the peer evaluation training and repeated practice equipped the intervention group to develop improved professional skills related to peer evaluation, compared with student autonomy and practice or training alone.

Limitations. The work reported in this paper does have some limitations to point out. First, the intervention group engaged in more peer evaluations throughout the semester versus the controls. Practicing behaviors or skills more often could have allowed more time for the intervention group to think about the quality of its feedback. Second, the instructors dedicated to grade the quality of the peer evaluations during the actual semester were different for the intervention and the control group. This inconsistency could have biased toward better quality instructor feedback by facilitating improvements in regard to the quality of the intervention groups’ evaluations. Third, we did not know the prior peer evaluation experiences.

Table A1. Student rubric for peer evaluation

| Peer-Grading Components | Constructive Comments: Indicate a Point Value and Elaborate |
|-------------------------|-------------------------------------------------------------|
| Content (20 points maximum): | Excellent (18–20); very good (15–17); fair (11–14); needs work (≤10) |
| Content of the presentation, including factual correctness and use of evidence to back assertions | Excellent (10); very good (8, 9); fair (6, 7); needs work (≤5) |
| Relevance (10 points maximum): | Excellent (10); very good (8, 9); fair (6, 7); needs work (≤5) |
| Integration of topic into the lecture materials, lecture goals, and overall course goals and competencies | Excellent (10); very good (8, 9); fair (6, 7); needs work (≤5) |
| Presentation (10 points maximum): | Excellent (10); very good (8, 9); fair (6, 7); needs work (≤5) |
| Clarity of the presentation, graphics, and text aided in understanding of the material, well-prepared, professional | Excellent (10); very good (8, 9); fair (6, 7); needs work (≤5) |
| Responsiveness (10 points maximum): | Excellent (10); very good (8, 9); fair (6, 7); needs work (≤5) |
| Responsiveness during question/answer session and moderating the online discussion | Excellent (10); very good (8, 9); fair (6, 7); needs work (≤5) |
| Additional comments and feedback | |

Advances in Physiology Education • doi:10.1152/advan.00114.2019 • http://advan.physiology.org
APPENDIX B: VALIDATED RUBRIC USED TO SCORE THE QUALITY OF STUDENT FEEDBACK

The rubric that the study investigators used to score the quality of the written student feedback is shown in Table B1.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

O.S.A. and D.B. conceived and designed research; O.S.A. and D.B. performed experiments; D.B. analyzed data; O.S.A., N.E.H., and D.B. interpreted results of experiments; D.B. prepared figures; O.S.A. and N.E.H. drafted manuscript; O.S.A., N.E.H., and D.B. edited and revised manuscript; O.S.A., N.E.H., and D.B. approved final version of manuscript.

REFERENCES

1. Aguilar SJ, Holman C, Fishman BJ. Game-inspired design: empirical evidence in support of gameful learning environments. Games Cult 13: 44–70, 2018. doi:10.1080/02602938.2018.1458211.
2. Allen SJ, Knight J. A method for collaboratively developing and validating a rubric. Int J Scholarly Teach Learn 3: 10, 2009. doi:10.20429/ijstl.2009.030210.
3. Bates DM, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4 (Preprint). ArXiv 1406.5823, 2014.
4. Borton K, Anderson OS. Metacognition gains in public health graduate students following in-class peer evaluation. Assess Eval High Educ 43: 1286–1293, 2018. doi:10.1080/02602938.2018.1458211.
5. Bridges D, Hisamatsu R, Anderson OS. Increasing student engagement within the core nutritional sciences curriculum: a gameful learning approach. Pedagog Heal Promot 5: 268–275, 2018. doi:10.1177/2373379918814022.

APPENDIX A: STUDENT RUBRIC FOR PEER EVALUATION

The grading rubric that students utilized to give their peers feedback according to different elements of the presentations is shown in Table A1.

Table A1. Validated rubric used to score the quality of student feedback

| Component          | Excellent (2 points)                                                                 | Good (1.5 points)                                                                 | Satisfactory (1 point)                                                                 | Needs Work (0.5 points)                                                                 | Poor (0 points)                                                                 |
|--------------------|--------------------------------------------------------------------------------------|----------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|
| Relevance          | All feedback aligns with both the assignment (in-class presentation) requirements and the peer evaluation rubric components. | Most feedback aligns with both the assignment requirements and the peer evaluation rubric components. | Some feedback aligned with both assignment requirements and the peer evaluation rubric components. | Some feedback aligned with either the assignment requirements and the peer evaluation rubric components, but not both | Feedback did not align with assignment requirements or the peer evaluation rubric components. |
| Specific           | All feedback is directed toward specific content or behavior.                        | Most feedback is directed toward specific content or behavior that can be altered. | Some feedback is directed toward specific content or a behavior that can be altered. | Feedback is general and does not target specific content or behaviors.                | Feedback does not target specific content or behaviors.                         |
| Realistic          | All feedback is within the scope of the assignment and could be altered by the student, given another opportunity. | Most feedback is within the scope of the assignment, which could be altered by the student, given another opportunity | Some feedback is within the scope of the assignment, which could be altered, given another opportunity. | Some feedback is within the scope of the assignment but could not all be altered, given another opportunity. | Feedback is too extreme to address within the scope of the assignment.          |
| Positivity         | Evaluator uses encouraging language and welcoming tone.                              | Evaluator mostly uses encouraging language and welcoming tone, but is not consistent. | Evaluator provides hints of encouragement and welcoming tone.                         | Evaluator lacks encouragement and welcoming tone, but is not discouraging.           | Evaluator uses discouraging tone.                                                |
| Professionalism    | If an employer reviewed this feedback, he/she would consider it complete and excellent. | If an employer reviewed this feedback, he/she would consider it complete and adequate. | If an employer reviewed this feedback, he/she would consider it complete but mostly adequate. | If an employer reviewed this feedback, he/she would consider it somewhat incomplete and mostly poor. | If an employer reviewed this feedback, he/she would consider it incomplete and unacceptable. |

Total points
6. Carr R, Palmer S, Hagel P. Active learning: the importance of developing a comprehensive measure. Active Learn High Educ 16: 173–186, 2015. doi:10.1177/1469787415589529.

7. Chadha D. A curriculum model for transferable skills development. Engl Educ 1: 19–24, 2006. doi:10.11120/ened.2006.01010019.

8. Chowdhury RR, Kalu G. Learning to give feedback in medical education. Obstet Gynaecol 6: 243–247, 2004. doi:10.1576/toag.6.4.243.27023.

9. Cranford Teague J. 8 Tips to Power-Up Your Classroom Presentations (Online). Edutopia. https://www.edutopia.org/blog/8-tips-classroom-presentation-jason-cranford-teague [17 Jan 2020].

10. Drummond I, Nixon I, Wiltshire J. Personal transferable skills in higher education: The problems of implementing good practice. Qual Assur Educ 6: 19–27, 1998. doi:10.1108/09684889810200359.

11. Hansen JG. Guiding principles for effective peer review. ELT J 59: 31–38, 2005. doi:10.1093/elt/cci004.

12. Hardavella G, Aamli-Gaagnat A, Saad N, Rousalova I, Sreter KB. How to give and receive feedback effectively. Breathe (Sheff) 13: 327–333, 2017. doi:10.1183/20734735.009917.

13. Holden JI, Kupperman J, Dorfman A, Saunders T, Pratt A, Mackay P. Gameful learning as a way of being. Int J Learn Technol 9: 181–201, 2014. doi:10.1504/IJLT.2014.064492.

14. Kiemen K, Grüschen A, Kunter M, Seidel T. Instructional and motivational classroom discourse and their relationship with teacher autonomy and competence support—findings from teacher professional development. Eur J Psychol Educ 33: 377–402, 2018. doi:10.1007/s10212-016-0324-7.

15. Klucervsek KM. Transferring skills from classroom to professional writing: student-faculty peer review as an extension of cognitive apprenticeship. J Scholarsh Teach Learn 16: 106–123, 2016. doi:10.14434/josoll.v16i6.20077.

16. Ko CT, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 15: 155–163, 2016 [Erratum in J Chiropr Med 16: 346, 2017]. doi:10.1016/j.jcm.2016.02.012.

17. Lam R. A peer review training workshop: coaching students to give and evaluate peer feedback. TESL Can J 27: 114, 2010. doi:10.18806/testl.v27i2.1052.

18. Lee H, Doh Y. A Study on the Relationship between Educational Achievement and Emotional Engagement in a Gameful Interface for Video Lecture Systems. 2012 International Symposium on Ubiquitous Virtual Reality, Adaejeon, South Korea, August 22–25, 2012. p. 34–37. doi:10.1109/ISUVR.2012.21.

19. Li L. Using game-based training to improve students’ assessment skills and intrinsic motivation in peer assessment. Innov Educ Teach Int 56: 423–433, 2019. doi:10.1080/14703297.2018.1511444.

20. Liou H-C, Peng Z-Y. Training effects on computer-mediated peer review. System 37: 514–525, 2009. doi:10.1016/j.system.2009.01.005.

21. Min H-T. Training students to become successful peer reviewers. System 33: 293–308, 2005. doi:10.1016/j.system.2004.11.003.

22. Min H-T. The effects of trained peer review on EFL students’ revision types and writing quality. J Second Lang Writ 15:118–141, 2006. doi:10.1016/j.jslw.2006.01.003.

23. Molloy E, Borrell-Carrió F, Epstein R. The impact of emotions in feedback. In: Feedback in Higher and Professional Education: Understanding It and Doing It Well, edited by Boud D, Molloy E. New York: Routledge, 2012. p. 60–81.

24. Nilson LB. Improving student peer feedback. Coll Teach 51: 34–38, 2003. doi:10.1080/07567550309596408.

25. Noziger AC, Naumburg EH, Davis BJ, Mooney CJ, Epstein RM. Impact of peer assessment on the professional development of medical students: a qualitative study. Acad Med 85:140–147, 2010. doi:10.1097/ACM.0b013e3181c47a5b.

26. Philippakos ZA. Giving feedback: preparing students for peer review and self-evaluation. Read Teach 71:13–22, 2017. doi:10.1002/trtr.1568.

27. Philippakos ZA, MacArthur CA. The effects of giving feedback on the persuasive writing of fourth- and fifth-grade students. Read Res Q 51:419–433, 2016. doi:10.1002/rrq.149.

28. Pope N. An examination of the use of peer rating for formative assessment in the context of the theory of consumption values. Assess Eval High Educ 26:235–246, 2001. doi:10.1080/02602930120052396.

29. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2019.

30. Rudenstine A, Schaaf S, Bacaalio D, Hakani S. Meeting Students Where They Are. Vienna, VA: iNACOL, 2018.

31. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 55:68–78, 2000. doi:10.1037/0003-066X.55.1.68.

32. Sargeant J, Mann K, Sinclair D, Van der Vleuten C, Metsemakers J. Understanding the influence of emotions and reflection upon multi-source feedback acceptance and use. Adv Health Sci Educ Theory Pract 13:275–288, 2008. doi:10.1007/s10459-006-9039-x.

33. Sarkar M, Overton T, Thompson C, Rayner G. Graduate employability: views of recent science graduates and employers. Int J Innov Sci Math Educ 24, 2016.

34. Schartel SA. Giving feedback—an integral part of education. Best Pract Res Clin Anaesthesiol 26:77–87, 2012. doi:10.1016/j.bpa.2012.02.003.

35. Shuijsmans DMA, Brand-Gruwel S, van Merriënoer JG. Peer assessment training in teacher education: effects on performance and perceptions. Assess Eval High Educ 27:443–454, 2002. doi:10.1080/026029302200009311.

36. van Zundert M, Shuijsmans D, van Merriënoer J. Effective peer assessment processes: Research findings and future directions. Learn Instr 20:270–279, 2010. doi:10.1016/j.learninstruc.2009.08.004.

37. Washington University in St. Louis, Center for Teaching and Learning. Using Peer Review to Help Students Improve Their Writing (Online). https://teachingcenter.wustl.edu/resources/writing-assignments-feedback/using-peer-review-to-help-students-improve-their-writing/ [17 Jan 2020].

38. Wen ML, Tsai C-C. Online peer assessment in an inservice science and mathematics teacher education course. Teach High Educ 13:55–67, 2008. doi:10.1080/13562510701794050.