3-2023

Science Assessments as a Learning Opportunity: Feedforward With Multiple Attempts

Emily Faulconer
Embry-Riddle Aeronautical University, faulcone@erau.edu

Science Assessments as a Learning Opportunity

Feedforward With Multiple Attempts

By Emily Faulconer and John Griffith

Feedback best practices support timely, high-quality feedback with application opportunity. Multiple attempts on assessments support learning gains. A learning management system can be used to automatically provide feedback for application on a future assessment attempt. Current research has not thoroughly investigated the student impacts or opinions on this combined strategy. In this study, students took a second attempt 46% of the time, scoring an average of 10.1% higher on their second attempt. More than 60% of students who failed their first attempt completed a second attempt. Students perceived the feedback as useful in preparing for their second attempt. Future research should include investigations of why some students do not make a second quiz attempt and in what ways feedback is used (or not). This study demonstrates the effectiveness of this feedforward with multiple attempts strategy in multiple introductory science courses taught fully online.

With the growth in online course offerings, much attention has been given to best practices in this modality. Learning management systems (LMS) such as Canvas and Blackboard are often used to deliver formative and summative assessments (Coates et al., 2005; Stodberg, 2012). Many of the LMS customizations of assessments provide guidance and support student engagement, which are pillars of exemplary assessments (Huba & Freed, 1999). For example, assessments can be programmed to automatically provide feedback immediately after submission, which is responsive and guides the students toward a stronger understanding of the gaps in their knowledge.

The literature supports timely feedback as a best practice (Gaytan & McEwen, 2007; Wiggins, 1993). In the online classroom, students reported that automatically generated feedback was more constructive than manual feedback (Bayerlein, 2014). By using the LMS to provide this feedback, instructors can provide the feedback consistently, using supportive language aligned with the assessment criteria. Assessment items that are conducive to automatic feedback include multiple-choice, true-or-false, fill-in-the-blank, and similar closed questions; automatic feedback on short answer, essay, or other open-response questions cannot be automatically graded by the LMS, so these formats are thus not suitable for automatic feedback as described in this study.

Crafting feedback that gives a student the opportunity to apply it—providing *feedforward*—is another best practice that makes feedback both engaging and learner centered (Hughes, 2011; Little et al., 2012; Wiggins, 2012). Feedforward allows students to demonstrate their mastery of learning objectives (Dulama & Ilovan, 2016; Goldsmith, 2010; Koen et al., 2012; Rodríguez-Gómez & Ibarra-Sáiz, 2015) and faculty to clarify their expectations (Baker & Zuvela, 2013). In online course assessments, these ideas can be implemented by embedding automatic feedback into the LMS while allowing students multiple attempts to answer a question.

The description of the use of multiple attempts in the literature is limited and has had varied parameters (e.g., testing time, presence of feedback, scoring of multiple attempts), making it challenging to draw conclusions (Orchard, 2016; Rhodes & Sarbaum, 2015; Yourstone et al., 2010). Some preliminary trends in the data are noted, though. The percentage of students who took advantage of multiple attempts varies, with studies reporting 36.5% in an in-person operations management course using online assessment (Orchard, 2016). In a study of chemistry lecture and lab assessments in asynchronous online courses, 74% of students who did not earn an A on the lecture assessment tried again, and 86% of students who did not earn an A on the laboratory assessment tried again (Faulconer et al., 2021). Some studies report that those who use multiple attempts may not outperform those who used only one (Faulconer et al., 2021; Orchard, 2016), and other studies reported gains for students who used multiple attempts (Rhodes & Sarbaum, 2015). However, because the students in these two groups earned similar final scores, this strategy may allow students to close the performance gap. Students may make use of a "throwaway attempt" to gain access to the feedback (Rhodes & Sarbaum, 2015; Yourstone et al., 2010), though some studies did not report evidence of this phenomenon (Faulconer et al., 2021).

The literature on multiple attempts combined with feedforward is significantly limited. In an online mathematics course, students were provided access to unlimited practice quizzes (ungraded) with automatically provided feedback, resulting in improved scores on summative assessment quizzes (Sancho-Vinuesa & Viladoms, 2012). However, a similar study in an online calculus course did not consistently show significant learning gains across terms (Sancho-Vinuesa et al., 2018). In our previous study, students demonstrated overall improvement in content mastery, as demonstrated through assessment grades (Faulconer et al., 2021). Student gains from this approach may be long-lasting, with significant gains reported on subsequent exams (Marden et al., 2013).

The lack of literature on the combination of these two strategies suggests a need for further research to demonstrate the effectiveness of combining formative feedback and multiple attempts on assessments. This work expands on our previous work that explored this construct in a single science discipline. The objective of this article is to demonstrate the benefits of these combined strategies in introductory courses in several science disciplines, including both learner outcomes and perspectives, thus demonstrating the applicability of the multiple attempts with feedforward scheme across subdisciplines. Our study explores the following hypotheses, reported as alternative hypotheses (H_a):

1. Students who do not earn an A on their initial attempt take advantage of the option to complete multiple attempts.
2. Students who do not earn a passing grade on their initial attempt take advantage of the option to complete multiple attempts.
3. Students' second attempt on the assessment outperforms their first attempt.
4. Students who take advantage of the option to complete multiple attempts outperform students who do not take advantage of this option.
5. The majority of students will report that feedback automatically provided after submitting their concept checks was useful.
6. The majority of students will report that they used the feedback provided on their first quiz attempt to prepare for their second quiz attempt.

Materials and methods

Participants

This study was performed at a medium-size private university. Due to the general trend in online education, the study's student population was nontraditional, with a higher average age, higher average level of employment, and higher rates of military affiliation than traditional students.

Courses selected for this study were introductory general education science courses that were available to science, technology, engineering, and mathematics (STEM) majors as well as nonmajors (see Table 1 for course and enrollment information). All enrollments originated from 9-week courses taught in the asynchronous online modality. Student performance data were obtained from the LMS between March 2018 and December 2019. Student perspectives on the usefulness of the feedback and

TABLE 1

Population and sample information.

Course	Enrollment (#)	Survey respondents (#)	Initial quiz attempts (#)	Second quiz attempts (#)
General Chemistry I	151	33	1,289	528
Introduction to Environmental Science	141	46	1,150	520
Science of Flight	119	31	1,072	569

TABLE 2

Example of feedback.

Topic	Question	Feedback
Significant figures	Which number below contains three significant figures?	You correctly identified that the zeroes to the left of the nonzero digits are never significant. You also recognized that zeroes in between nonzero numbers are always significant. Keep in mind that zeroes to the right of nonzero numbers are significant if there is a decimal present.
Subatomic particles and atomic models	How should this diagram be changed to properly represent the lithium-8 isotope? [diagram not pictured]	You recognized that electrons are outside of the nucleus and are not changed between various isotopes of an element. Elements are arranged within the periodic table by increasing atomic number. Because atomic number defines an element, isotopes of an element have the same atomic number (which represents the number of protons). However, isotopes of an element have different atomic masses (which represents the number of protons and neutrons). Which of these two subatomic particle varies in isotopes?
Rock cycle	A geologist discovers an intrusion of igneous rock that cuts through four layers of sedimentary rock. What layer of rock is the oldest?	Sedimentary rock forms from settling and compaction. Would older sediments appear near the top or bottom of the rock layer?

self-reported behaviors regarding feedback use were obtained from end-of-course evaluations between May 2019 and December 2019, resulting in a response rate ranging from 63.3% to 69.4%. Performance data were obtained through nonprobability sampling, and student perception data were obtained through self-selected sampling (Sterba & Foster, 2008).

To protect participants' identity, all data were aggregated with no individual identifiers. Because the literature supports both automatic feedback and multiple attempts, no control group was utilized. Instead, the student use of the multiple attempts with feedforward assessment design in a course was explored. This work was reviewed by the Institutional Review Board and deemed exempt.

Assessment design

There were nine summative assessments in each course, with weight ranging from 25% to 40% of the overall grade, meaning that the weight of each assessment in calculating the overall grade ranged from 2.78% to 4.44%. There was no penalty for only using one attempt. The highest score was awarded as the final assessment grade. In each course in this study, assessments were administered through the LMS, pulling questions from pools aligned with learning objectives. This means that each assessment attempt was unique for each student and each attempt, although the complexity of the problems and the content area alignment were controlled through the use of the objective-aligned pools. This approach prevented question familiarity and addressed the issue of student tendency to select the same wrong response on a second attempt (Feinberg et al., 2015). Assessment questions were closed (e.g., multiple choice, multiple answer, true or false) and presented one at a time, with no open responses for any assessments. The question pools were written at a level of Bloom's taxonomy that aligned with the correlated learning objective. However, some higher-level learning objectives also had additional question pools written at the lower level.

Because the multiple-attempts scenario is likely to increase student time on task, the assessment was timed to reduce the likelihood of an extensive amount of time added to student workload. Each attempt was limited to 1 hour, though students could stop, save their work, and resume later. Each assessment began with a brief statement to inform students of the option for multiple attempts, when to expect feedback, and how to best use the feedback. Similar language to communicate this assessment design to students was included in the course syllabus and course announcements.

The LMS automatically graded the assessments, with feedback provided once, immediately upon completion of the attempt. Correct answers were not provided by the LMS or within the feedback. With examples provided in Table 2, the feedback programmed into the LMS was designed on the principles of high-quality feedback: specific, actionable, timely, and supportive (Bayerlein, 2014; Huba & Freed, 1999). Within the LMS, instructors also provided feedback to students after the assessment due date. With the approach, the feedback in this course aligns with the well-supported philosophy that feedback is a mechanism for enhancing learning (Hattie & Timperley, 2007).

Survey data

Student perceptions regarding the automatically provided feedback were collected by adding custom questions to the institutionally standardized end-of-course evaluations administered online. Using a 5-point Likert scale, respondents were asked to state their level of agreement with the following statements:

- The feedback automatically provided after submitting my quiz was useful.
- I used the feedback provided on my first quiz attempt to prepare for my second quiz attempt.

The surveys were completed anonymously (with no individual identifiers, including IP address), with data aggregated. Survey data were used to evaluate the last two hypotheses. In those cases, the "strongly agree" and "agree" were combined into an "agree" category and "neutral," "disagree," and "strongly disagree" responses were categorized as "disagree" (Gay et al., 2006). The categories were combined to allow for effective evaluation of Hypotheses 5 and 6 and to ensure assumptions of the chi-squared statistic (independent observations and cell sizes equal to or greater than 5) were not violated.

Statistical analysis

A total of 3,511 initial and 1,617 second attempt scores were used to evaluate the first four hypotheses in this study. All data analysis was performed using StatCrunch on the internet or Statdisk (Triola, 2013). The first two hypotheses regarding (i) the number of students who scored below an A and (ii) those who did not achieve a passing score in the first attempt were tested using chi-squared ($\alpha = 0.025$) due to the nominal nature of the data.

The third hypothesis, regarding whether students outperformed their first quiz score on their second attempt, was evaluated using a one-tailed paired-samples *t*-test ($\alpha = 0.025$). The fourth hypothesis, which concerned if students who took advantage of the multiple attempts outperformed students who did not, was evaluated with a one-tailed two-sample *t*-test ($\alpha = 0.025$). Finally, survey responses regarding automatic feedback after concept checks (Hypothesis 5) and feedback after the first attempt on quizzes (Hypothesis 6) were evaluated using chi-squared ($\alpha = 0.025$) due to the nominal nature of the data. The alpha settings reflect a Bonferroni-adjusted alpha (from 0.05) due to the relationships between hypothesis pairs that were grouped into "families."

Hypotheses 1 and 2 used some of the same data to evaluate both hypotheses. A similar Bonferroni correction was made for the "family" of Hypotheses 3 and 4, as well as the "family" of Hypotheses 5 and 6. These Bonferroni corrections (using a lower alpha) were designed to avoid type 1 errors. Results were then evaluated using the appropriate effect size test (Gould & Ryan, 2012).

Results and discussion

All alternative hypotheses in this study were supported by student data, validating the multiple attempts with feedforward assessment design. Each hypothesis is discussed in detail in the following sections.

Student utilization of the option to complete multiple attempts: Feedforward assessment design

Students' motivation to complete a second attempt might vary based on their score on their first attempt. The first and second hypotheses address the tendency to utilize a second attempt based on the score of the first attempt (< A and < D, respectively). Students who do not earn an A on their initial attempt take advantage of the option to complete multiple attempts. Among the students with the 2,863 initial scores that fell below 90% (an A), students elected to retake the quiz 1,524 times in the courses examined. The chi-squared test of good fit analysis yielded significant results with a small phi effect size ($X^2 = 11.95$, $p < 0.001$, $\varphi = 0.064$).

A higher percentage (more than 53%) of students elected to retake the quiz. Chemistry and environmental science students who achieved a passing score below 90% were not more likely to take additional attempts. However, a significant majority of students in the science of flight course who achieved a score below 90% took advantage of a second attempt ($X^2 = 44.1$, $p < 0.001$, $\varphi = 0.234$; small phi effect size).

Students who do not earn a passing grade on their initial attempt take advantage of the multiple attempts

(Hypothesis 2). Of students whose 655 initial scores fell below passing, students elected to retake the quiz 464 times. The chi-squared test of good fit yielded significant results, with a medium phi effect size (X^2 = 113.8, $p < 0.001$, $\varphi = 0.417$; medium phi effect size). Of the students who earned a failing score on their first attempt, more than 60% elected to retake the quiz.

These trends from the full data set were consistent across the individual disciplines studied, with each science discipline showing significant results:

- Chemistry: $X^2 = 11.84$, $p < 0.001$, $\varphi = 0.197$ (small phi effect size)
- Environmental science: $X^2 = 52.6$, $p < 0.001$, $\varphi = 0.515$ (large phi effect size)
- Science of flight: $X^2 = 80.53$, p < 0.001, $\varphi = 0.725$ (large phi effect size)

These data are consistent with existing literature that reports a utilization of multiple attempts (in multiple science and nonscience disciplines) ranging from 35% to 95% (Faulconer et al., 2021; Orchard, 2016; Stewart et al., 2014).

Student impact due to the multiple attempts with feedforward assessment design

To validate the learner outcome benefits of multiple attempts that were hypothesized based on reported benefits of feedforward and multiple attempts, this study explored how the multiple attempts with feedforward assessment design influenced student grades. Completing a second attempt requires an additional time investment (Faulconer et al., 2021). We wanted to investigate whether the time investment to complete a second attempt was worth students' time.

Students' second attempt on the assessment outperforms their first attempt (Hypothesis 3). Of the 1,617 of the times students retook the quizzes overall, on 1,183 attempts (more than 73%), they achieved an average of 10.1 points higher on the second attempt (out of 100 points total). The right-tailed paired-samples *t*-test yielded significant findings and medium Crohn's *d* effect size $t = 23.575$, $p < 0.001$, $d = 0.586$. As implied, 27% of students had the same score or a lower score on their second attempt. Three different science courses were evaluated in this research. However, all showed similar findings, ranging between a 9% and 11.8% average improvement on the second attempt. This is consistent with data from our previous studies (Faulconer et al., 2021). The assessment design that uses question pools suggests that this improvement in the grade is due to authentic content knowledge gains.

Students who take advantage of the multiple attempts outperform students who do not take advantage of the multiple attempts (Hypothesis 4). In this study, 46% of students who elected to retake the quizzes earned higher quiz scores. The difference in final assessment scores (the better of the two attempts vs. the result of a single attempt) was 79.9% vs 79.1%. A right-tailed two-sample *t*-test did not yield significant results using the Bonferroni corrected alpha of 0.025 and small Hedges' *g* effect size; $t(N = 3510) = 1.66$, $p = .0488$, $g = 0.055$. In this study, students who invested the time to complete a second attempt fared better than students who did not. The difference in this study was not significant because it was still relatively small (less than 1%). For this reason, these results tend to be consistent with the previous literature that reports no significant difference (Faulconer et al., 2021; Orchard, 2016).

Although it is possible that students could use a "throwaway" attempt to gain access to the feedback, there was no clear evidence of this in the data from this study. If present, this approach would skew the data in favor of a positive impact of a second attempt. However, claims in this area would require bold assumptions regarding student motivations, which are not justified without qualitative data. What we did see, however, was the potential "abandonment" of the second attempt, where students would perform very poorly. While no claims can be made without qualitative evidence, a potentially abandoned attempt may have multiple unanswered questions and a short time investment. If second attempts were abandoned, it would diminish the positive impacts of good-faith second attempts reported in our study.

Student perspectives on the multiple attempts with feedforward assessment design

Two survey questions were used to evaluate student perceptions of the multiple attempts with feedforward assessment design. The survey inquired about (i) the usability of feedback provided through the feedforward concept and (ii) whether the students used the feedback prior to making a second quiz attempt.

The majority of students reported that feedback automatically provided after submitting their assessment was useful (Hypothesis 5). Of the 110 students who responded, 92 agreed or strongly agreed (83.6%) that the feedback was useful, yielding a significant chi-squared test of good fit result and large phi effect size ($X^2 = 49.78$, $p < 0.001$, $\varphi = 0.673$).

The majority of students reported

that they used the feedback provided on their first attempt to prepare for their second attempt (Hypothesis 6). Of the 109 students who responded, 92 agreed or strongly agreed (84.4%) that they used the feedback to prepare for the second quiz attempt, yielding a significant chi-squared test of good fit result and large phi effect size ($X^2 = 51.61$, $p < 0.001$, $\varphi = 0.688$).

More than half of students who did not earn an A on their first attempt used a second attempt. It can be assumed that at least some students retook the quizzes without remedial study of the topic area being assessed. However, nearly three-quarters of the second attempts showed improvement. This improvement likely drove the strong positive response for the survey. However, without qualitative data, it is not possible to draw further conclusions regarding if and how students remediated or why some students did not find the feedback useful or chose not to apply it.

Limitations

There are several limitations in this study. The primary limitation was the lack of a control group in validating the multiple attempts with feedforward assessment design model. However, given the efficacy of the separate constructs and the previously published data by the authors that demonstrated efficacy, there may be ethical concerns with establishing a control group.

It is challenging to control all moderating variables in a field experiment. By their own admission through the survey, not all students used second attempts. It is not clear if those who did not use the feedback were also those who did not use a second attempt. As mentioned earlier, "throwaway" and "abandoned" attempts can influence the data.

This study used a nontraditional student population. The average age was 34, and students had a higher level of employment than traditional students. Additionally, the population was approximately 50% active duty and reserve military and 30% military affiliated. Military student demographics in higher education are similar to nontraditional students (Ford & Vignare, 2015). Like nontraditional students, military students tend to complete their coursework online (Ford & Vignare, 2015). Demographics may influence results within certain subcategories. This study is designed to look at high-level trends, protecting participants' confidentiality and privacy through anonymous data collection, to prevent the analysis of differences in performance and perspectives among subgroups of the population. Comparison of this study to those performed using traditional students or an in-person student population may by restricted due to demographics as a potential moderating factor in student use of multiple attempts or application of feedback before completing a second attempt, as well as moderating student perspectives of usefulness. Future work should seek confidential rather than anonymous data so that demographic-moderating variables can be explored.

All courses were delivered asynchronously online in a 9-week format. Results may differ between this study and any replication that uses a traditional student population following a typical 16-week term schedule.

Approximately 27% of the time, students did not score higher on their second attempt quiz on a particular topic. Investigation as to the possible causes of this result was outside the scope of the current study.

Conclusions

This study examined the impact of the multiple attempt with feedforward assessment design. Key conclusions are as follows:

1. If students are offered multiple attempts on summative assessments in various science disciplines, they tend to take advantage of the opportunity.
2. Students who use multiple attempts on summative assessments in various science disciplines tend to demonstrate stronger content mastery on their second attempt.
3. Students in various science disciplines find feedforward useful and apply it to multiple attempts.
4. Student utilization of, impact of, and perspectives on the multiple attempts with feedforward assessment design is consistent across several science disciplines.

A multiple attempt with feedforward assessment design can be designed within the LMS, turning the first attempt's feedback into formative feedback because students are offered the opportunity to apply the feedback to improve the gaps in their knowledge. Although there is a time investment required of faculty to prepare this feedback scheme, the results support this effort, as it has been shown to improve student assessment scores across multiple science disciplines. Once the assessments are designed, however, the time investment to continue the construct is minimal, and instructors can have more time to provide personalized, detailed feedback.

Future work is needed to further validate this pedagogical choice. Specifically, a qualitative study could explore student reasons for abstaining from multiple attempts and why some students did not apply the feedback.

Issues of self-efficacy or experience with college learning environments could be contributing factors. Additionally, a qualitative study could more accurately describe "throw-away" and "abandoned" attempts. It would also be interesting to explore how the level of cognitive learning according to Bloom's taxonomy influences results.

References

Baker, D. J., & Zuvela, D. (2013). Feedforward strategies in the first-year experience of online and distributed learning environments. *Assessment & Evaluation in Higher Education*, *38*(6), 687–697. https://doi.org/10.1080/02602938.2012.691153

Bayerlein, L. (2014). Students' feedback preferences: How do students react to timely and automatically generated assessment feedback? *Assessment & Evaluation in Higher Education*, *39*(8), 916–931. https://doi.org/10.1080/02602938.2013.870531

Coates, H., James, R., & Baldwin, G. (2005). A critical examination of the effects of learning management systems on university teaching and learning. *Tertiary Education and Management*, *11*(1), 19–36. https://doi.org/10.1007/s11233-004-3567-9

Dulama, M. E., & Ilovan, O. (2016). How powerful is feedforward in university education? A case study in Romanian geography education on increasing learning efficiency. *Journal of Educational Sciences: Theory & Practice*, *16*(3), 827–848. https://doi.org/10.12738/estp.2016.3.0392

Faulconer, E. K., Griffith, J. C., & Frank, H. (2021). If at first you do not succeed: Student behavior when provided feedforward with multiple trials for online summative assessments. *Teaching in Higher Education*, *26*(4), 586–601. https://doi.org/10.1080/13562517.2019.1664454

Feinberg, R. A., Raymond, M. R., & Haist, S. A. (2015). Repeat testing effects on credentialing exams: Are repeaters misinformed or uninformed? *Educational Measurement: Issues and Practice*, *24*(1), 34–39. https://doi.org/10.1111/emip.12059

Ford, K., & Vignare, K. (2015). The evolving military learner population: A review of the literature. *Online Learning*, *19*(1), 7–30.

Gay, L. R., Mills, G. E., & Airasian, P. W. (2006). *Educational research: Competencies for analysis and applications* (8th ed.). Pearson Education.

Gaytan, J., & McEwen, B. C. (2007). Effective online instructional and assessment strategies. *American Journal of Distance Education*, *21*(3), 117–132. https://doi.org/10.1080/08923640701341653

Goldsmith, M. (2010). *Try feedforward instead of feedback.* Art From Ashes. https://www.artfromashes.org/wp-content/uploads/2017/03/FeedForward-Article.pdf

Gould, R. N., & Ryan, C. N. (2012). *Introductory statistics* (1st ed.). Pearson.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Huba, M. E., & Freed, J. E. (1999). *Learner-centered assessments on college campuses: Shifting the focus from teaching to learning.* Allyn & Bacon.

Hughes, G. (2011). Towards a personal best: A case for introducing ipsative assessment in higher education. *Studies in Higher Education*, *36*(3), 353–367. https://doi.org/10.1080/03075079.2010.486859

Koen, K., Bitzer, E. M., & Beets, P. A. D. (2012). Feedback or feedforward? A case study in one higher education classroom. *Journal of Social Sciences*, *32*(3), 231–242. https://doi.org/10.1080/09718923.2012.11893068

Little, J. L., Bjork, E. L., Bjork, R. A., & Angello, G. (2012). Multiple-choice tests exonerated, at least of some charges: Fostering test-induced learning and avoiding test-induced forgetting. *Journal of Psychological Science*, *23*(22), 1337–1344. https://doi.org/10.1177/0956797612443370

Marden, N. Y., Ulman, L. G., Wilson, F. S., & Velan, G. M. (2013). Online feedback assessments in physiology: Effects on students' learning experiences and outcomes. *Advances in Physiology Education*, *37*(2), 192–200. https://doi.org/10.1152/advan.00092.2012

Orchard, R. K. (2016). Multiple attempts for online assessments in an operations management course: An exploration. *Journal of Education for Business*, *91*(8), 427–433. https://doi.org/10.1080/08832323.2016.1256262

Rhodes, M. T., & Sarbaum, J. K. (2015). Online homework management systems: Should we allow multiple attempts? *American Economist*, *60*(2), 120–131. https://doi.org/10.1177/056943451506000203

Rodríguez-Gómez, G., & Ibarra-Sáiz, M. S. (2015). Assessment as learning and empowerment: Towards sustainable learning in higher education. In M. Peris-Ortiz & J. Merigo Lindahl (Eds.), *Sustainable learning in higher education* (pp. 1–20). Springer. https://doi.org/10.1007/978-3-319-10804-9_1

Sancho-Vinuesa, T., Masià, R., Fuertes-Alpiste, M., & Molas-Castells, N. (2018). Exploring the effectiveness of continuous activity with automatic feedback in online calculus. *Computer Applications in Engineering Education*, *26*(1), 62–74. https://doi.org/10.1002/cae.21861

Sancho-Vinuesa, T., & Viladoms, N.

E. (2012). A proposal for formative assessment with automatic feedback on an online mathematics subject. *International Journal of Educational Technology in Higher Education*, *9*(2), 241–260. https://doi.org/10.7238/rusc.v9i2.1285

Sterba, S. K., & Foster, E. M. (2008). Self-selected sample. In P. J. Lavrakas (Ed.), *Encyclopedia of survey research methods* (pp. 806–808). SAGE Publications. https://dx.doi.org/10.4135/9781412963947.n525

Stewart, D., Panus, P., Hagemeier, N., Thigpen, J., & Brooks, L. (2014). Pharmacy student self-testing as a predictor of examination performance. *American Journal of Pharmaceutical Education*, *78*(2), Article 32. https://doi.org/10.5688/ajpe78232

Stodberg, U. (2012). A research review of e-assessment. *Assessment & Evaluation in Higher Education*, *37*(5), 591–604. https://doi.org/10.1080/02602938.2011.557496

Triola, M. (2013). *Statdisk* [computer software]. Pearson Education. https://www.statdisk.com/

Wiggins, G. (2012). Seven keys to effective feedback. In M. Scherer (Ed.), *On formative assessment: Readings from educational leadership* (pp. 24–35). Association for Supervision and Curriculum Development.

Wiggins, G. P. (1993). *Assessing student performance: Exploring the purpose and limits of testing*. Jossey-Bass.

Yourstone, S. A., Kraye, H. S., & Albaum, G. (2010). Online quantitative-based assignments—are more attempts better for learning? *Decision Sciences Journal of Innovative Education*, *8*(2), 347–351. https://doi.org/10.1111/j.1540-4609.2010.00260.x

Emily Faulconer *(*faulcone@erau.edu*) and* ***John Griffith*** *are associate professors, both in the Department of Mathematics, Science, and Technology at Embry-Riddle Aeronautical University in Daytona Beach, Florida.*