Online feedback assessments in physiology: effects on students’ learning experiences and outcomes

Nicole Y. Marden, Lesley G. Ulman, Fiona S. Wilson, and Gary M. Velan

Departments of Physiology, School of Medical Sciences, The University of New South Wales, Sydney, New South Wales, Australia; and Department of Pathology, School of Medical Sciences, The University of New South Wales, Sydney, New South Wales, Australia

Submitted 28 June 2012; accepted in final form 4 March 2013

Marden NY, Ulman LG, Wilson FS, Velan GM. Online feedback assessments in physiology: effects on students’ learning experiences and outcomes. Adv Physiol Educ 37: 192–200, 2013; doi:10.1152/advan.00092.2012.—Online formative assessments have become increasingly popular; however, formal evidence supporting their educational benefits is limited. This study investigated the impact of online feedback quizzes on the learning experiences and outcomes of undergraduate students enrolled in an introductory physiology course. Four quiz models were tested, which differed in the amount of credit available, the number of attempts permitted, and whether the quizzes were invigilated or unsupervised, timed or untimed, or open or closed book. All quizzes were composed of multiple-choice questions and provided individualized feedback. Summative end-of-course examination marks were analyzed with respect to performance in quizzes and were also compared with examination performance in the year before the quizzes were introduced. Online surveys were conducted to gather students’ perceptions regarding the quizzes. The vast majority of students perceived online quizzes as a valuable learning tool. For all quiz models tested, there was a significant relationship between performance in quizzes and end-of-course examination scores. Importantly, students who performed poorly in quizzes were more likely to fail the examination, suggesting that formative online quizzes may be a useful tool to identify students in need of assistance. Of the four quiz models, only one quiz model was associated with a significant increase in mean examination performance. This model had the strongest formative focus, allowing multiple unsupervised and untimed attempts. This study suggests that the format of online formative assessments is critical in achieving the desired impact on student learning. Specifically, such assessments are most effective when they are low stakes.

online assessment; formative assessment; summative assessment; feedback; learning outcomes

However, several studies (7, 8) have shown that when formative assessments are optional with no course credit assigned, there is limited use of them by students. Thus, it is becoming more popular to assign a small amount of credit to encourage students’ engagement with formative assessments (8, 13).

Paper-based formative assessments have a number of limitations: in particular, providing individualized feedback is time consuming and may not be feasible for large classes (21). Online formative assessments are a time-effective alternative, allowing staff to provide timely feedback to large groups of students without creating an overwhelming workload. Online formative assessments offer several potential advantages to students: they offer students more flexibility in terms of time and place and they provide immediate feedback and can provide links to learning resources thereby motivating further study (14, 21). While online formative assessments are becoming an increasingly popular supplement to traditional summative examinations in higher education, formal evidence to support their educational benefits is relatively limited and often contradictory; some studies (1, 6, 14) have shown that formative assessments do not improve the learning outcomes of students, whereas others (2, 3, 8, 12, 20) have shown that they do have a positive effect on students’ performance in summative examinations.

The University of New South Wales offers introductory undergraduate physiology courses in the second year of study. These courses are offered in two consecutive sessions, the first of which covers excitable cell physiology, muscle physiology, cardiovascular physiology, blood, and neurophysiology. The second session course covers endocrine and reproductive physiology, respiratory physiology, renal physiology, and gastrointestinal physiology. These courses are undertaken by students enrolled in a wide range of degrees, including Science, Medical Science, Optometry, Exercise Physiology, and Engineering degrees. Given the diverse student cohort, there can be considerable variation in university entrance marks, background knowledge, and level of interest in the subject matter among our students. Student enrollments are very large: the session 1 course typically attracts between 420 and 500 students and session 2 enrollments are typically between 320 and 350 students. Although introductory, these physiology courses are comprehensive in scope, and a large proportion of students find them extremely challenging. The large student numbers and diverse student cohort pose a significant challenge to teaching staff in terms of motivating and engaging students, providing feedback on performance, and identifying and assisting students in need of remediation.

Address for reprint requests and other correspondence: N. Y. Marden, Dept. of Physiology, School of Medical Sciences, The Univ. of New South Wales, Sydney 2052, NSW, Australia (e-mail: n.marden@unsw.edu.au).

1043-4046/13 Copyright © 2013 The American Physiological Society
Beginning in 2009, we introduced online formative assessments in both session 1 and session 2 introductory physiology courses in an effort to provide students with individual feedback on their learning progress, to enhance motivation and engagement with the course material, and, ultimately, to improve learning outcomes. In previous years, we had only used traditional summative assessments in our course, with students being graded on a midsession examination and an end-of-session examination (consisting of a combination of multiple-choice questions and short-answer written responses) as well as a series of prelaboratory quizzes. The quizzes were structured to provide instant feedback to students, thereby enabling them to identify and correct misconceptions earlier in the learning process. Several different formats of online quizzes were used over a period of three consecutive years (2009–2011). The formats differed in a number of aspects, including the amount of course credit available per quiz, whether the quizzes were supervised or unsupervised, the number of attempts permitted per quiz, whether the quizzes were timed, and whether the students were permitted to use study aids (such as lecture notes and textbooks) while attempting the quizzes.

This study evaluated the impact of the different models of online feedback quizzes that we implemented on the learning experiences and learning outcomes of the cohorts of students enrolled in our session 1 and session 2 introductory physiology courses between 2009 and 2011.

METHODS

Online quizzes comprising multiple-choice questions were authored and delivered using Questionmark Perception software (Questionmark) and were designed such that after students submitted their responses (or at the time limit), they received immediate, individualized feedback. We chose to use multiple-choice questions as these are the style of questions delivered in our summative examinations, and questions were constructed to be of a similar nature and standard to those used in summative examinations. Figure 1 shows an example of a typical multiple-choice question used in the online quizzes and the corresponding feedback students received for an incorrect response. Students accessed the quizzes via the university’s learning management system using secure individual logins.

Several different presentations of online quizzes were applied throughout the 3-yr period of this study beginning in session 1 of 2009. Changes between one quiz model and the next were motivated by the overall aim of improving student learning outcomes by stimulating meaningful and consistent engagement with course material and were based on ongoing analysis of student outcomes combined with consideration of feedback on the quizzes gathered from students and staff. The four different quiz models implemented with key features are shown in Table 1.

In session 1 of 2009, quiz models 1 and 2 were applied and differed only in whether or not the assessments were supervised. Three online feedback quizzes were offered during the 12-wk course in weeks 5, 9, and 12. To encourage student participation and preparation, performance in each quiz contributed 5% toward the overall course mark. Each quiz consisted of 10 multiple-choice questions with a 20-min time limit, focusing on a specific section of the course. Quizzes 1 and 2 (model 1) were made available for students to attempt unsupervised in their own time within a 48-h timeframe. Students were instructed to attempt these quizzes independently and without the use of lecture notes and study aids. Quiz 3 (model 2) was conducted under supervision during laboratory class time. The change to an invigilated setting for quiz 3 was in response to reports that students had been
attempting quizzes 1 and 2 in groups, with the aid of their lecture notes.

In session 2 of 2009, quiz model 3 was implemented. In this model, we decided to retain the invigilated format of model 2 that was used for quiz 3 in session 1; however, only two quizzes were scheduled due to timetabling constraints (one quiz in week 6 and one quiz in week 12). The number of multiple-choice questions in each quiz was increased from 10 to 15. To keep the overall credit value of these quizzes the same as in the session 1 course, each quiz contributed 7.5% toward the overall course mark. To encourage students to prepare thorough study notes, the quizzes in model 3 were changed to an open-book format. We retained quiz model 3 in both session 1 and session 2 courses in 2010.

In sessions 1 and 2 of 2011, quiz model 4 was implemented. This model was significantly different from the previous models in a number of key aspects, and these considerable changes were motivated by the lack of effect on student outcomes that had been observed with the earlier quiz models over the previous 2 yr. In this model, five untimed quizzes were offered per session, with each quiz consisting of multiple-choice questions relating to one topic within the course. Quizzes were made available online for students to attempt within a few days after the conclusion of each course topic, and students attempted the quizzes unsupervised in their own time. Each quiz was accessible online for a period of 1 wk, and students were permitted to attempt the quiz as many times as they wanted within this period. Quizzes were set up such that the order in which the questions were presented was randomly changed for each quiz attempt, as were the order of the five alternatives within each question. Each quiz contributed 2% toward the students' overall course grade and was awarded provided the student received a minimum score of 90% for the quiz. After the conclusion of the 1-wk open period for each quiz, data were collected and marks were awarded, and each quiz was then made available again for unlimited review until the summative examinations. The number of multiple-choice questions in each quiz varied depending on the number of lectures devoted to that particular area in the lecture series. For example, reproductive physiology was covered in 3 lectures and the online feedback quiz on reproductive physiology consisted of 16 multiple-choice questions, whereas there were 9 lectures devoted to the topic of cardiovascular physiology, and the cardiovascular physiology online feedback quiz consisted of 47 multiple-choice questions. This model also differed to the previous quiz models tested in that the multiple-choice questions examining lecture material in the summative examinations at the end of session were drawn from the bank of questions used in the online feedback quizzes throughout the session. This was done to maintain strong student participation and engagement with this quiz model despite the lower-stakes format.

With the exception of session 1 of 2010, a voluntary, anonymous online evaluation survey was conducted after the final quiz in each session to gather students’ perceptions regarding the quizzes. There was no survey conducted at the conclusion of the quizzes in session 1 of 2010 due to the absence of a key staff member at that time.

Our study was approved by the Medical and Community Human Research Ethics Advisory Panel of The University of New South Wales (reference nos. 2009-7-06 and 2011-7-07). Consent for participation in the feedback component of the study was implied by the response to the anonymous online survey. Students completed online feedback quizzes and the end-of-session examination as part of their learning in the introductory physiology courses.

Summative end-of-session examination marks were analyzed with respect to performance in quizzes and were also compared with end-of-session examination marks achieved by students completing the courses in 2008, in which the quizzes were not offered. All statistical analyses were performed using IBM SPSS (version 20). Comparisons of student survey responses between cohorts were performed using Kruskal-Wallis tests. To compare mean examination marks between cohorts, ANOVAs were used, followed by Dunnett’s t-tests as appropriate. To compare quiz scores between groups of students within each cohort who passed or failed end-of-course examinations, Student’s t-tests were used. Relationships between performance in online formative assessments and end-of-course examinations were determined using correlation and linear regression analyses.

The withdrawal rates from the courses were consistent across all sessions studied, ranging from 0% to 1.8%.

RESULTS

Data gathered from the student evaluation surveys offered at the end of each session showed that most students’ perceptions of the online feedback quizzes were very favorable (Fig. 2). Student response rates for these online evaluation surveys were generally very high: 86% for session 1 of 2009, 95% for session 2 of 2009, 87% for session 2 of 2010, 80% for session 1 of 2011, and 70% for session 2 of 2011. The slight decrease in response rates in 2011 might reflect the change from the earlier models used in 2009 and 2010, in which the quizzes (and subsequent evaluation surveys) were conducted in an invigilated setting where staff could encourage student participation in the evaluation surveys, to quiz model 4, in which the quizzes and evaluation surveys were taken by students in their own time. Despite the changing features of the different quiz models applied, evaluation surveys consistently showed that the vast majority of students found the quizzes challenging and valuable for guiding study and providing feedback. This was consistent across all sessions of this study regardless of the quiz model applied; no significant differences were found in median percent agreement for each survey question between cohorts (Fig. 2).

In the open-ended section of the evaluation survey, many students commented that the best features of the online quizzes

---

**Survey Questions**

Fig. 2. Students’ responses to online feedback surveys. Data are percentages of positive student responses (i.e., strongly agree or agree) to each of five Likert-scale questions asked in an anonymous online evaluation survey offered at the end of each session from 2009 to 2011 [note: due to the absence of a key staff member, student evaluation of online feedback quizzes in session 1 (S1) of 2010 was not conducted]. Median percent agreement for each question was compared between cohorts using Kruskal-Wallis tests, and no significant differences were found. Numbers in bars are exact percentages; 393 of 455 students who could have potentially accessed the survey responded in S1 of 2009 (86%), 330 of 347 students who could have potentially accessed the survey responded in S2 of 2010 (87%), 387 of 482 students who could have potentially accessed the survey responded in S1 of 2011 (80%), and 276 of 396 students who could have potentially accessed the survey responded in S2 of 2011 (70%).
were that they provided immediate feedback to help correct misconceptions and that they stimulated more frequent revision of the course material throughout the session, thus making preparation for the final examination easier. The open-ended section of the survey proved to be quite useful in gathering insights into students’ thoughts regarding some of the variables of the different quiz models applied over the course of this study (for example, open vs. closed book, invigilated vs. unsupervised, and higher vs. lower credit). Some examples of students’ comments about the quizzes are included below.

Session 1 of 2009 (quiz models 1 and 2). The following are student comments regarding session 1 of 2009 (quiz models 1 and 2):

Instant feedback was important because it allowed you to instantly correct something in your mind that you had wrong, rather than go on thinking that until you got your marks back, or in some cases like the midsession, only get marks and don’t actually know which questions you got right or wrong.

The best features were the exposure to exam style questions before the actual exam. This allowed for more thorough review of the content and a realization of weaknesses in my learning.

They broke up the various topics within the course and made sure that continual studying was occurring. It has made me learn and revise the course content instead of leaving it until the end of the semester.

Session 2 2009 and 2010 (quiz model 3). The following are student comments regarding session 2 of 2009 and 2010 (quiz model 3):

Being able to bring your own notes in kept me up to date with lecture notes and forced me to study more throughout semester and be organized with notes throughout semester.

Open book exam reduced the exam stress a lot. That way I felt more confident and was able to use my knowledge more efficiently.

The open book nature of the quiz made it less stressful, more enjoyable, and a valuable learning tool to clear up misconceptions, focus on problem areas as well as give students an indication of where they needed to focus their learning.

Sessions 1 and 2 2011 (quiz model 4). The following are student comments regarding sessions 1 and 2 of 2011 (quiz model 4):

This was an exceptional learning tool. It helped to guide my learning and push me further to research and learn about each concept individually, not just so I could get the marks but to obtain understanding. These quizzes have been the best learning device I have used (that has been supplied by the university) in my 3 years studying at UNSW. It was fantastic.

The challenging nature of the questions encouraged pursuing different approaches to understanding the content tested. The quizzes could be done from the comfort of the student’s home. The shuffling of questions and answers within each question with each attempt ensured the input of more effort on the student’s part and made what was tested “stick” better.

The quizzes are a great way for revising the material given in lectures and a way of making sure that you are keeping up to speed with the lecture topics presented. The fact that multiple goes were allowed with an unlimited amount of time meant that you could think about the questions without feeling rushed to make stupid mistakes.

Extensive quizzes on subject matter will always help me understand and learn the processes better and I think they should be adopted in all schools. The unlimited attempts makes it an even better study aid as the quizzes are attempted without an air of pressure (usually resulting in people finding ways to cheat) and the small grade incentive makes sure most students at least study in a small way.

Excellent revision tool. I would read over the lectures first, then attempt them first time under exam conditions–then redo them with materials.

The impact of introducing online feedback quizzes on student learning outcomes was evaluated by comparing the mean summative end of session examination mark of students taking session 1 and session 2 courses in 2009, 2010, and 2011 (in which various models of the online quizzes were offered) with those taking the course in 2008 (in which the online quizzes were not offered). There were no significant differences in mean end-of-session examination marks between students taking the session 1 course in 2009 (55.47 ± 0.78, n = 455), in which quiz models 1 and 2 were offered, and those taking the corresponding course in session 1 of 2008 (57.61 ± 0.65, n = 467), in which there were no quizzes (Fig. 3A). Similarly, there were no significant differences in mean summative end-of-session examination marks between students taking the session 2 course in 2009 (55.67 ± 0.74, n = 347), in which quiz model 3 was implemented, and those taking the session 2 course in 2008 (54.42 ± 0.81, n = 313), in which there were no quizzes (Fig. 3B). Quiz model 3 was applied again in both session 1 and session 2 courses in 2010, and no significant differences in mean summative end-of-session examination marks were observed between the 2008 or 2009 student cohorts and the 2010 student cohorts for either session 1 (55.38 ± 0.68, n = 419; Fig. 3A) or session 2 (53.02 ± 0.78, n = 321; Fig. 3B). In sessions 1 and 2 of 2011, quiz model 4 was applied, and the mean summative examination marks for students in both sessions (session 1 of 2011: 61.31 ± 0.60, n = 482 (Fig. 3A) and session 2 of 2011: 59.76 ± 0.70, n = 396 (Fig. 3B)) were significantly higher than those in the corresponding courses in 2008, 2009, and 2010.

To determine whether performance in online quizzes may be useful in predicting performance in summative examinations, correlations and regression analyses of students’ online quiz scores and summative end-of-session examination marks were performed. There was a significant linear relationship between performance in quizzes and performance in the end-of-session examination, which was evident for all quiz models tested; mean performance in online feedback quizzes accounted for 24.4%, 36.4%, and 7.8% of the variance in end-of-session examination marks in session 1 of 2009, 2010, and 2011, respectively (P < 0.001; Table 2) and for 24%, 23.3% and 6.4% of the variance in end-of-session examination marks in session 2 of 2009, 2010, and 2011, respectively (P < 0.001; Table 2).
Correlation analysis between scores in individual quizzes and summative end-of-session examination marks was also performed, and significant relationships were observed for each of the quizzes in both sessions in 2009 and 2010 (Table 2). In quiz model 4, which was applied in sessions 1 and 2 in 2011, there was a significant change in the scoring of the quizzes, with students receiving a mark of either 0 or 2 for each quiz depending on whether or not they achieved a minimum score of 90% for the quiz. This change in scoring meant that it was not meaningful to perform correlation analysis between scores in individual quizzes and summative end-of-session examination marks in 2011.

Interestingly, performance in invigilated quizzes (models 2 and 3) accounted for a larger proportion of variance in end-of-session examination marks than performance in quizzes that were not invigilated (models 1 and 4). In session 1 of 2009, quiz 3 (model 2), which was conducted under supervision, accounted for a larger proportion of variance in end-of-session examination marks (16.5%) than performance in quizzes 1 and 2 (model 1), which were not invigilated (13.1% and 7.8%, respectively; Table 2). In model 3, which was applied in session 2 of 2009 and sessions 1 and 2 of 2010, quizzes were conducted under supervision but differed from model 2 in that they were open book. The fact that these quizzes were open book did not diminish their correlation with performance in the end-of-session examination (Table 2).

The relationship between performance in online feedback quizzes and performance in summative examinations was further investigated by dividing students into two categories, those who passed and those who failed the end-of-session examination, and mean scores were compared for both the individual quizzes and overall online quiz marks for those two groups. Again, due to the change in scoring for quiz model 4, this analysis was only performed for mean rather than individual quizzes in 2011. As shown in Table 3, students who performed poorly in the online feedback quizzes were more likely to fail the end-of-session examination. This was the case for each quiz analyzed in all courses (P < 0.001; Table 3) as well as for the mean of all quiz scores in all courses (P = 0.002 for session 2 of 2011 and P < 0.001 for all other courses; Table 3); the difference in mean online feedback quiz scores between students who passed or failed the end-of-session examination was 2.39, 2.07, and 1.94, 1.91, and 0.62 out of 10 for session 2 of 2009, 2010, and 2011, respectively (P < 0.001), and 1.22, 1.19, and 0.62 out of 10 for session 3 of 2009, 2010 (P < 0.001), and 2011 (P = 0.002), respectively (Table 3).

**DISCUSSION**

This study investigated the impact of introducing online feedback quizzes on the learning experiences and outcomes of several large cohorts of students undertaking introductory physiology courses at the University of New South Wales between 2009 and 2011. Several different models of online feedback quizzes were implemented throughout the study period, with differing impact on learning outcomes. Cohorts of students undertaking the courses in which quizzes were offered in the format of models 1–3 did not demonstrate any significant overall improvement in learning outcomes, as measured by performance in the summative end-of-session examinations. In contrast, the implementation of quizzes in the format of model 4 in sessions 1 and 2 of 2011 was associated with a significant improvement in mean summative end-of-session examination scores. Data from student evaluation surveys indicated that the vast majority of students welcomed the introduction of online feedback quizzes and perceived them to be valuable in terms of stimulating engagement with the course material and guiding study. The online quizzes were found to be predictive of performance in summative end-of-session examinations; students who performed poorly in the online feedback quizzes were more likely to fail the end-of-session examination. This was evident for all quiz models tested, although the relationship between performance in quizzes and...
The varying impact of different quiz models on student learning outcomes seen in our study is not surprising, given the contrasting results of similar studies reported in the literature. Peat and Franklin (14) investigated the use of both paper-based and online assessment resources by a large group of first-year undergraduate biology students and found that offering these formative assessments had no significant impact on final learning outcomes, despite being perceived as a useful learning tool by the students. In fact, that study reported that a greater proportion of students who failed the course had made use of the formative assessment resources compared with the more successful students. Similarly, a study of students in a functional anatomy course conducted by Urtel et al. (19) found that the introduction of online quizzes increased engagement of students in the classroom but did not result in an improvement in academic performance.

Nevertheless, many other studies (2–4, 8, 12, 20) have demonstrated a positive impact of online formative assessments on student learning. Dobson (4) found that the introduction of online formative assessments into an undergraduate exercise physiology course was associated with a significant improvement in summative examination results. Similarly, studies conducted by Kibble (8), Olsen and McDonald (12), and Velan et al. (20) in undergraduate medical or dentistry students have all shown that participation in online formative assessments is associated with a significantly better performance in summative examinations.

These contrasting examples in the literature raise the question as to why such variability regarding the impact of online formative assessments on student learning exists and whether there are certain elements within the design or implementation of online formative assessments that are critically important in achieving the desired impact on student learning outcomes. In the present study, changes in the key features of each quiz model were based on analysis of student outcomes, together with staff and student reflections on the quizzes, and only the final of the four models tested was associated with a significant improvement in student learning outcomes.

After reviewing the lack of effect of online quizzes on end-of-session examination performance in the first session of our study (session 1 of 2009), we hypothesized that this might have been related to undesirable behavior elicited by the format of model 1 that was used for two of the three quizzes offered in that course. Students were advised to prepare thoroughly for the quizzes and then attempt them individually and without the use of their lecture notes and textbooks. While some students complied with the instructions, we became aware that many others were attempting the quizzes in groups without prior preparation and instead simply relying on study aids to answer the questions. Reports of such undesirable behavior prompted us to change to an invigilated setting for the final quiz in session 1 of 2009 (model 2), and we retained the supervised format for model 3 used in session 2 of 2009 and sessions 1 and 2 of 2010. Despite the invigilated format of the quizzes in models 2 and 3, there was still no significant improvement in students’ performance in the end-of-session examination associated with the application of these models.

While models 2 and 3 were similar in that they were both invigilated, a key point of difference was that the quizzes in model 3 were open book, in which students were allowed to use study resources (lecture notes and textbooks) to assist them with the quiz questions. The shift to an open-book format for model 3 was intended to encourage students to prepare and develop their own study notes. Comments provided by students in the open-ended section of the evaluation surveys suggest that this aim was indeed achieved, with many students commenting that the open-book format encouraged them to be more organised and prepare good summary notes from the lectures. An additional benefit commented upon was that the
open-book format reduced exam stress and made the quizzes more enjoyable.

While offering the online quizzes in an open-book format has the apparent advantage of stimulating organisation and preparation by students, the lack of effect of this model on performance in the end-of-session examinations might be related to students using their study resources to answer the quiz questions; this is likely to have improved quiz scores, and students may have interpreted high quiz scores as an indication that they had mastered the content when this may have not been the case.

The quizzes that we introduced provided feedback to students but were not purely formative; while it was not mandatory for students to complete the quizzes to pass the course overall, for models 1–3 in particular, there was a significant amount of course credit attached to the quizzes (5% per quiz in models 1 and 2 and 7.5% per quiz in model 3), which resulted in almost all students participating in the online quizzes. While attaching significant course credit to the quizzes was clearly successful in achieving strong student participation, it might have led to students perceiving that the stakes for those assessments were high, which may have detracted from the learning benefits associated with formative assessment. Indeed, this aligns with evidence that in order for formative assessments to be effective, students must perceive them to be relatively low stakes (12) and nonthreatening (18).

With these considerations in mind, and given the lack of effect of models 1–3, significant changes led to the development of model 4, which was intended to lower the stakes and give the quizzes a more formative focus. The number of quizzes offered per session was increased, and the credit attached to them was decreased. Students attempted the quizzes unsupervised and with no time limit. In contrast to all previous models, students were allowed to attempt the quiz as many times as they wanted within a 1-wk open access period.

<table>
<thead>
<tr>
<th>Session 1 of 2009</th>
<th>Mean Online Feedback Quiz Marks</th>
<th>Number of students</th>
<th>Mean score</th>
<th>Mean difference</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online feedback quiz 1 (quiz model 1)</td>
<td>Pass</td>
<td>281</td>
<td>7.55/10</td>
<td>1.72</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Fail</td>
<td>164</td>
<td>5.83/10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online feedback quiz 2 (quiz model 1)</td>
<td>Pass</td>
<td>278</td>
<td>8.76/10</td>
<td>1.00</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Fail</td>
<td>162</td>
<td>7.76/10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online feedback quiz 3 (quiz model 2)</td>
<td>Pass</td>
<td>275</td>
<td>6.68/10</td>
<td>1.83</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Fail</td>
<td>156</td>
<td>4.85/10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean online feedback quizzes</td>
<td>Pass</td>
<td>283</td>
<td>11.31/15</td>
<td>2.39</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Fail</td>
<td>167</td>
<td>8.92/15</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Session 2 of 2009</th>
<th>Mean Online Feedback Quiz Marks</th>
<th>Number of students</th>
<th>Mean score</th>
<th>Mean difference</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online feedback quiz 1 (quiz model 3)</td>
<td>Pass</td>
<td>217</td>
<td>14.07/15</td>
<td>1.45</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Fail</td>
<td>129</td>
<td>12.63/15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online feedback quiz 2 (quiz model 3)</td>
<td>Pass</td>
<td>217</td>
<td>11.50/15</td>
<td>2.48</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Fail</td>
<td>129</td>
<td>9.02/15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean online feedback quizzes</td>
<td>Pass</td>
<td>217</td>
<td>12.78/15</td>
<td>1.94</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Fail</td>
<td>129</td>
<td>10.84/15</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Session 1 of 2010</th>
<th>Mean Online Feedback Quiz Marks</th>
<th>Number of students</th>
<th>Mean score</th>
<th>Mean difference</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online feedback quiz 1 (quiz model 3)</td>
<td>Pass</td>
<td>265</td>
<td>9.87/15</td>
<td>1.97</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Fail</td>
<td>147</td>
<td>7.90/15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online feedback quiz 2 (quiz model 3)</td>
<td>Pass</td>
<td>264</td>
<td>10.84/15</td>
<td>1.91</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Fail</td>
<td>148</td>
<td>8.93/15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean online feedback quizzes</td>
<td>Pass</td>
<td>266</td>
<td>10.33/15</td>
<td>2.07</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Fail</td>
<td>151</td>
<td>8.27/15</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Session 2 of 2010</th>
<th>Mean Online Feedback Quiz Marks</th>
<th>Number of students</th>
<th>Mean score</th>
<th>Mean difference</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online feedback quiz 1 (quiz model 3)</td>
<td>Pass</td>
<td>186</td>
<td>13.69/15</td>
<td>1.60</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Fail</td>
<td>131</td>
<td>12.09/15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online feedback quiz 2 (quiz model 3)</td>
<td>Pass</td>
<td>184</td>
<td>11.34/15</td>
<td>2.01</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Fail</td>
<td>130</td>
<td>9.33/15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean online feedback quizzes</td>
<td>Pass</td>
<td>186</td>
<td>12.45/15</td>
<td>1.91</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Fail</td>
<td>134</td>
<td>10.55/15</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Session 1 of 2011</th>
<th>Mean Online Feedback Quiz Marks</th>
<th>Number of students</th>
<th>Mean score</th>
<th>Mean difference</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total online feedback quizzes (quiz model 4)</td>
<td>Pass</td>
<td>396</td>
<td>9.43/10</td>
<td>1.12</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Fail</td>
<td>86</td>
<td>8.30/10</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Session 2 of 2011</th>
<th>Mean Online Feedback Quiz Marks</th>
<th>Number of students</th>
<th>Mean score</th>
<th>Mean difference</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total online feedback quizzes (quiz model 4)</td>
<td>Pass</td>
<td>298</td>
<td>9.27/10</td>
<td>0.62</td>
<td>0.002</td>
</tr>
<tr>
<td>Fail</td>
<td>98</td>
<td>8.65/10</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Statistical analysis of the relationship between performance in online feedback quizzes and failure in the end-of-session examination in each course was performed by Student’s t-tests.
learning tool that can be used to test and enhance understanding. More emphasis was also placed on the students’ approach to quizzing: students were advised to prepare for the quizzes and attempt them under exam conditions, at least in the first instance, to provide a realistic appraisal of their performance and understanding of the content being tested. Students were further encouraged to make a note of items answered incorrectly and to revise any areas of uncertainty before attempting the quiz again. Data gathered from the evaluation surveys suggest that many students were indeed approaching and using the quizzes in the way we had intended. Allowing unlimited untimed attempts appears to have been critical in facilitating this approach, with students commenting that these features allowed them the freedom to test their knowledge and enhance their understanding without the primary concern of scoring credit or fear of being penalised. At the same time, the allocation of a small amount of course credit once an acceptable level of competency was achieved appears to have provided sufficient incentive to participate in the quizzes, with student participation in this model being similar to those previous models in which higher course credit was attached to individual quizzes. The fact that questions in the summative examinations were drawn from the question bank used for the quizzes is also likely to have provided an incentive to participate.

This latter model of online quizzes appears to have had a positive impact on student learning outcomes; mean summative end-of-session examination marks for students taking the course in the sessions in which quiz model 4 was applied (sessions 1 and 2 of 2011) were significantly higher than those taking previous iterations of the course in which quizzes were not offered or were offered in the higher-stakes formats of models 1–3.

It is possible that these cohorts of students performed significantly better in the summative end-of-session examination not because of the enhanced learning offered by the model 4 quizzes but because they simply memorized the questions and answers from the quizzes and retrieved these memorized facts for the end-of-session examination. However, we believe this is less likely for the following reasons.

1. The results of our study align with a growing body of research indicating that taking a test on studied material enhances learning and later retention compared with simply undertaking additional study of the material (10, 16, 17). This phenomenon, known as the testing effect or test-enhanced learning, is most powerful when, as in our final quiz model, testing occurs frequently and students receive feedback on the tests (10, 17). In some of these studies (10, 16), as in ours, summative examinations contained identical questions or very similar questions to those previously presented to students in quizzes, but another study (10) has shown that the testing effect is still evident when an entirely different set of questions or even a different format (e.g., short-response instead of multiple-choice questions) are used for the final summative examinations compared with the sessional quizzes.

2. The online quizzes contained a large number of questions on each topic area compared with the number of questions used in the summative examinations; thus, students would have needed to memorize a considerable bank of questions to answer the relatively small number appearing again in the summative examination.

3. For each quiz attempt, both the order of questions and the order of alternatives were randomized and were different to that presented in the summative examination, thereby discouraging a direct memorization approach.

4. Qualitative data collected from the open-ended feedback sections of the student evaluation surveys strongly suggest that students were approaching the online quizzes in model 4 in the way we had intended (i.e., as a learning tool to enhance understanding rather than a purely grade-focused assessment exercise), thus supporting the notion that the improvement in summative end-of-session examinations by this cohort of students is the result of the enhanced learning stimulated by this quiz format.

Another major finding of this study was that the online feedback quizzes we implemented demonstrate predictive validity. For all quiz models tested, we observed a significant correlation between performance in quizzes and performance in the end-of-session examination. Those students who failed the end-of-session examination performed significantly worse in the online quizzes than those who passed the end-of-session examination. While overall course outcome is influenced by course credit obtained in other assessments, including a midsession examination and practical quizzes, a high proportion of students who failed the summative end-of-session examination failed the course overall. Hence, the online quizzes might have an important diagnostic role, by indicating which students are at risk of failure, and therefore might benefit from early intervention, such as academic counselling and remediation.

In summary, students undertaking introductory physiology courses at the University of New South Wales have welcomed the introduction of online feedback assessments, with the majority of students perceiving them to be a useful learning tool. For all quiz models tested in this study, performance in online feedback quizzes was significantly correlated with performance in summative end-of-session examinations; thus, online quizzes are good predictors of final exam performance and could be used to target students in need of remediation and assistance. The format of online formative assessments appears to be critically important in achieving the desired impact on student learning outcomes; of the four different online quiz models tested, only one quiz model was associated with a significant improvement in summative end-of-session examination performance. This was the most formative of the models tested, with features specifically designed to shift students’ perceptions and use of the online quizzes from a credit scoring exercise to that of a learning tool to explore and enhance understanding. Thus, our study supports the notion that in order for online formative assessments to be effective, they must be perceived by students to be relatively low stakes and allow exploration of existing and expected knowledge in a nonthreatening environment.

ACKNOWLEDGMENTS

The authors thank Prof. Gary Housley for suggestions on improving the manuscript.

DISCLOSURES

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

Author contributions: N.Y.M., L.G.U., and F.W. conception and design of research; N.Y.M., L.G.U., and F.W. performed experiments; N.Y.M., L.G.U., and G.V. interpreted results of experiments; N.Y.M. and L.G.U. prepared figures; N.Y.M. and L.G.U. drafted manuscript; N.Y.M. and G.V. edited and revised manuscript; N.Y.M. and G.V. approved final version of manuscript; G.V. analyzed data.

REFERENCES