The use of formative online quizzes to enhance class preparation and scores on summative exams

John L. Dobson
Department of Applied Physiology and Kinesiology, University of Florida, Gainesville, Florida

Submitted 11 July 2008; accepted in final form 24 September 2008

Dobson JL. The use of formative online quizzes to enhance class preparation and scores on summative exams. Adv Physiol Educ 32: 297–302, 2008; doi:10.1152/advan.90162.2008.—Online quizzes were introduced into an undergraduate Exercise Physiology course to encourage students to read ahead and think critically about the course material before coming to class. The purpose of the study was to determine if the use of the online quizzes was associated with improvements in summative exam scores and if the online quizzes were valid predictors of summative exam performance. A retrospective analysis was performed on the course scores from three different groups of Exercise Physiology students. Students in group 1 completed the original version of the course, those in group 2 completed an updated version of the course that included more rigorous exam questions, and those in group 3 completed the same updated version of the course but with the addition of 10 required online quizzes. Results showed that the overall mean summative exam score from group 3 was significantly higher than that from group 2 (81.79 ± 8.26 and 78.72 ± 9.61, respectively). A significant positive correlation (r = 0.50) was also found between individual mean online quiz scores and individual mean exam scores for those students in group 3. It was concluded that the formative online quizzes did enhance summative exam performance and that the online quizzes were valid predictors of exam performance.

formative assessments; exercise physiology; undergraduate

TO MAXIMIZE STUDENT ACHIEVEMENT, teachers need to provide students with plenty of effective feedback (12) using a range of different assessments (9, 20). To this end, it is becoming increasingly common for teachers to use formative assessments to supplement their traditional summative exams. Whereas the intent of a summative exam is to show students how they did at the end of an educational process, a formative assessment provides feedback about how they are doing along the way (19). Formative assessments can be made using any type of instrument that evaluates understanding, but they are most effective when they are of the same type of instrument as the summative assessments they supplement (19). For example, if a summative exam is comprised of true/false and multiple-choice questions, then a complementary formative assessment (e.g., quiz) should also be composed of true/false and multiple-choice questions. Nevertheless, formative assessments, unlike summative exams, must be viewed by students to be relatively “low stake” (15) or “nonthreatening” (14, 19) evaluations for them to be effective. It is therefore common for formative assessments to be offered on a volunteer basis and for no course credit (15, 18, 19). Still, students will only use formative assessments if they also view them to be worthwhile (17), and so it is not uncommon for teachers to assign some small level of course credit to them (14, 16). Consequently, formative assessments are not necessarily characterized as evaluations that have no penalty. Rather, they tend to be broadly defined as relatively unintimidating assessments that are used to provide rapid feedback to students early in the evaluation process (1, 19).

The feedback from formative assessments allows students to better understand the teacher’s expectations, and it allows the teacher to modify the course work to better address the students’ needs (17, 19). Indeed, the evidence suggests that formative assessments may benefit both students and teachers in numerous ways. For example, formative assessments have been shown to encourage students to be better prepared for class (8), to have less test anxiety (21), and to have a more positive attitude about their class performance (7, 13, 21). Furthermore, although some studies (3, 11, 171 18) have found that the use of formative assessments does not enhance overall learning outcomes, other studies (2, 5, 6, 14, 15, 21) have indicated that they may be used to enhance summative exam scores.

In light of the many benefits discussed above, formative assessments were introduced into one of the fall 2005 undergraduate Exercise Physiology courses at the University of Florida. The intent of the formative assessments was to encourage students to read ahead in the course lecture notes and textbook so that they would be better prepared for class meetings. The assessments took the form of online quizzes because they could provide instantaneous feedback to the students without creating additional grading responsibilities. Now that the online quizzes have been used in several subsequent Exercise Physiology courses, a retrospective analysis was performed to determine if 1) the use of the online quizzes was associated with improvements in summative exam scores and 2) if the online quizzes were valid predictors of summative exam performance.

METHODS

Background. The Department of Applied Physiology and Kinesiology (APK) at the University of Florida has undergraduate majors in exercise physiology, fitness-wellness, and athletic training. APK majors, like the majority of undergraduates at the University of Florida, are highly intelligent and competitive students. For example, the typical enrolling student scored at least a 1250 on her or his SATs and had a high school grade point average of at least 3.75.

All APK undergraduates are required to take Exercise Physiology in either their sophomore or junior year. This course, which is an extension of its prerequisite Applied Human Physiology, focuses on the metabolic, endocrine, muscular, cardiovascular, and respiratory responses to both acute and chronic exercise. A typical enrollment in Exercise Physiology is ~40 students/class, and there are generally at least 1 or 2 classes offered every fall, spring, and summer semester.
The first time I taught the Exercise Physiology course was during the fall of 2001. Although I had a decent idea of the material I had wanted to cover and how I had wanted to present it during that first offering, many of the elements of that course were entirely new to me at the time. Of particular importance, it was the first time I had created and incorporated both extensive PowerPoint presentations and intensive exam questions into a course. Although those original lectures and exams were acceptable initial efforts, it became clear to me, as I taught the class again in both the subsequent spring and summer semesters, that I needed to revise the course materials to make them more effective. Therefore, I spent the latter portion of the summer of 2002 rethinking my course lecture notes and revising my PowerPoint slides so that I could present the material in a more targeted and more cohesive way. I also spent a great deal of time revising and updating the course exam questions so that they would be more challenging for the students. My goal, when making those exam question modifications, was to make it much more difficult for students to guess the correct answer to the questions. Examples of those exam question modifications are presented in Tables 1 and 2. Using the questions shown in Table 1 as an example, the students should have known that each of the first three answer choices (answers A, B, and C) in both of the questions was true. With the original version of the question, however, the students really only needed to know that two of those three choices were true to “guess” the correct answer (answer D). With the modified version of the question, only those students that knew that answers A, B, and C were all true were likely to select the correct answer (answer D).

The inclusion of the revisions mentioned above into my fall 2002 and spring 2003 Exercise Physiology courses resulted in significant increases in both my student evaluations and my own subjective ratings of the quality of the course. Yet, those revisions, as well as my accumulating experience with the course, also exposed a new set of concerns with the class. The first concern was that the students were being evaluated solely on the basis of four summative exams. Therefore, unlike in most other APK courses, the students in Exercise Physiology completed no other assignments and had no additional means of earning course credit, other than the summative exams. Once I had modified those exams, and more importantly as the students’ scores declined due the enhanced difficulty of the questions, the students began to complain that there were no alternative means of earning points in the class.

A second, and more significant, growing concern that I had been having was that my students were not preparing for each class meeting by reading ahead and thinking critically about the material as I had requested. As I frequently explained to my students at that time, the purpose of that requirement was to give us the opportunity to then use our class meetings to discuss, rather than to introduce, the course material. Unfortunately, the students were not following the requirement, and so, instead of participating in a dialogue about the material, they were all too often unengaged and passive throughout the class meetings.

My concerns about the lack of student preparation and my exclusive reliance on summative exams in my Exercise Physiology course stayed with me during the following 2 years (2003–2005) when I left the University of Florida. It was during that time away that I decided that I would begin using online quizzes to both supplement the summative exams and to encourage students to better prepare for class meetings. When I returned to the University of Florida in the fall of 2005, I first introduced these formative online quizzes into my Exercise Physiology course.

Online quizzes. The online quizzes were administered using the University of Florida’s WebCT system. This system was chosen to administer the quizzes because it automatically and rapidly both graded the assessments and provided feedback to students. The system also automatically generated reports that showed both how every student performed on every quiz and how all students performed on each question within a quiz.

A total of 10 quizzes were offered each semester, 1 quiz during roughly each academic week in which a summative exam was not scheduled. Each quiz was worth 1% of the total course points. Each quiz consisted of 10 true/false and/or multiple-choice questions that pertained to a specific section of the course textbook and/or lecture notes. Examples of typical quiz questions are shown in Tables 3 and 4. Each quiz was programmed to open for students on a specific day and time (typically on a weekend day) and then close exactly 48 h later (typically on the first weekend day) and then close exactly 48 h later (typically on the first weekend day) and then close exactly 48 h later (typically on the first weekend day) and then close exactly 48 h later (typically on the first weekend day).
class meeting day of week). The purpose of this window of time was to encourage students to read the appropriate course material just before (i.e., within 48 h of when) they would be discussing it in class during the subsequent week.

Within the 48-h period in which each quiz was available, students were also given exactly 15 min to complete the quiz after they had begun it. If a student exceeded the 15-min time limit, the WebCT system would automatically close the quiz and prevent the student from answering further. The time limit was imposed for two related reasons. First, although students were allowed to consult their textbook and notes throughout each quiz, the short 15-min duration ensured that students could only be successful on the quiz if they prepared in advance. The second reason was that previous research (4) indicated that timed electronic quizzes eventually better enhance exam performance more than untimed quizzes.

**Experimental groups.** There were three different groups of students in this investigation. Students in group 1 were those that took the Exercise Physiology course in its initial form and with the original exam questions. This group included all of my Exercise Physiology students from the fall 2001, spring 2002, and summer 2002 semesters. Students in group 2 took the course after the content and exam questions had been modified to be more rigorous. This group included all of my Exercise Physiology students from the fall 2002 and spring 2003 semesters. Students in group 3 took the course after the formative online quizzes had been added. This group included all of my Exercise Physiology students from the fall 2005, fall 2006, spring 2007, and fall 2007 semesters.

**Data analysis.** All experimental procedures were approved by the University of Florida Institutional Review Board. Data analysis was conducted retrospectively for all semesters during the summer of 2008. Data were taken from the nine undergraduate Exercise Physiology courses taught at the University of Florida between the years of 2001 and 2007. To maintain student anonymity, each individual student’s name was replaced with an alphanumeric code. Each student’s 4 summative exam scores were averaged, as were the 10 online lecture quiz scores for each student in group 3. Statistical comparisons between the overall mean summative exam scores for students in groups 1, 2, and 3 were made using ANOVA with post hoc t-tests. An additional analysis was performed to determine the effect size ($\eta^2$) pertaining to comparison between overall mean summative exam scores in groups 2 and 3. Finally, for those students in group 3, a correlation was made between individual mean online quiz scores and individual mean summative exam scores. Statistical significance was set at $P < 0.05$. Individual student performance on examinations and online lecture quizzes was reported as a percentage, and data are expressed as means $\pm$ SD.

**RESULTS**

**Student and course data.** Assessment data were taken from a total of 417 students. There were 132 students in group 1, 97 students in group 2, and 188 students in group 3. Specific numbers of students and mean summative exam scores from each of the nine different Exercise Physiology courses are shown in Table 5.
Discussion

Summary of main findings. The primary purpose of this investigation was to determine how the use of online formative quizzes affected summative exam performance in undergraduate Exercise Physiology classes. The main conclusions of this study were that 1) online quiz use was associated with improvements in summative exam scores and student learning and 2) online quizzes were valid predictors of summative exam performance.

Comparison of group summative exam scores. The major distinction between groups 1 and 2 was that the students in the latter group took Exercise Physiology after the summative exam questions had been modified with the goal of making them more rigorous. The significantly reduced overall mean summative exam score from group 2 versus that from group 1 indicates that the modification was a success and that the exam questions had become more challenging.

Despite the fact that nearly 2 years passed between the last class section of group 2 and the first class section of group 3, all of the key elements of the course remained the same throughout that time and for both groups. That is, the course material, lecture notes, PowerPoint slides, and summative exam questions used by groups 2 and 3 were all identical. Nevertheless, students in group 3 had significantly greater overall mean summative exam scores than did students in group 2. Although the effect size pertaining to the difference in those group scores can be considered small to modest, it was comparable with the effect size reported by a similar previous investigation (15). Furthermore, the mean summative exam score for group 2 was exactly in the middle of the range that corresponded to a grade of C+ (i.e., within the range of 77.5–79.9), whereas the mean score for group 3 was well above the minimum needed to earn a B (i.e., above 80.0). Therefore, the difference in mean summative exam scores was large enough to be very meaningful to the students.

There are several possible explanations for the significant difference in summative exam scores between groups 2 and 3. First, it is possible that the instructor did a relatively better job helping the students in group 3. This is, however, extremely unlikely since every element of the course, including the exact progression of material, the format of the class, the types of interactions, and the specific descriptions given for each and every concept, were very close to identical. A second possibility is that the students in group 3 possessed different characteristics that gave them some advantage on the summative exams. This is also not likely to have had a significant influence on the results because each group contained a large number of students, and both the enrollment criteria and course prerequisites were the same for all of the students in the investigation. A third possibility is that the students in group 2 somehow unfairly informed those in group 3 as to exactly what their summative exam questions would be asking. It is, however, doubtful that this type of cheating occurred because the students in the two groups probably did not overlap in time enough to really interact with one another. More importantly, all copies of the exams were always collected and accounted for promptly after each use, and the students were not allowed to transcribe or make notes regarding the questions. Therefore, it is highly unlikely that any copies of, or specific information about, the summative questions or exams were even available to those in group 3.

The final, and much more likely, explanation for the significantly higher summative exam scores in group 3 is that those students benefited from the one resource that was not available to those in group 2: the online formative quizzes. It therefore appears that the online quizzes successfully allowed the students in group 3 to better understand the teacher’s expectations and the teacher to modify the course work to better address those students’ needs. There are also two pieces of evidence that indicate the quizzes benefited the students in group 3 by enhancing their preparation for each class meeting. First, there was a subjective, but very real, sense that the students in group 3 were more familiar with, and better able to discuss, the material than were those in group 2. Second, and more importantly, the overall mean online quiz score for students in group 3 was high (85.61 ± 9.95). Since there is very little chance that those students simply guessed correctly 85% of the time, they must have been well prepared for those quizzes. It is possible that they were well prepared because they shared the correct answers with one another. Still, the cost-to-benefit ratio of such an action was likely not favorable for those students because they were well aware that to have shared answers would have meant that they had committed an honor violation that carried serious consequences. It is more likely that they were well prepared for those quizzes because they first learned the quiz material, as they were asked to do. The majority of the quiz material had not been covered in other prerequisite classes, so the students had to have initiated their exposure to the material by reading the course textbook and studying the lecture notes. Since the quizzes preceded the lectures in which the corresponding material was discussed, the students likely read ahead.
and studied the lecture notes before, and were therefore more prepared for, the lectures and class discussions. It is reasonable to conclude that this relatively more effective class preparation, in turn, helped students in group 3 perform better on their summative exams.

Interestingly, the overall mean summative exam scores for groups 1 and 3 were both not only greater than the overall mean score from group 2 but also statistically similar to one another. That would seem to indicate that the enhanced difficulty of the modified summative exam questions was effectively counterbalanced by the benefit of the online quizzes on summative exam performance. It is nice to know that the course had evolved such that the most recent students, likely through better class preparation, were just as successful with more rigorous summative exams as were the previous students with their easier exams. Better still, it is reasonable to conclude that the clearer and more concise course materials, more rigorous exams, higher course expectations, formative quizzes, and greater summative exam performance all combined to make the final version of the course (group 3) the most effective at enhancing student learning.

One final interesting outcome that pertained to the summative exam scores was the difference in the mean scores from the spring 2007 and fall 2007 classes. Both of those classes were in group 3, which means that both classes completed the online quizzes and both took the same exams. Nevertheless, the spring 2007 mean summative exam score was slightly lower than the overall group 3 mean, the fall 2007 mean exam score were slightly higher than the overall group 3 mean, and there was a significant difference in the mean scores of the two individual classes. Since all of the elements of both classes were identical, and because no other classes differed significantly from the others in their group, it seems reasonable that the significant difference in the spring and fall 2007 class scores may be explained by differences in the characteristics of the relatively small cohort of students that were enrolled in each class.

It could be argued that the online quizzes used in this investigation were not formative assessments because they were not offered on a volunteer basis and they were associated with course credit. Nonetheless, other refereed studies (14, 16) have also classified assessments associated with course credit as formative. Furthermore, the course credit associated with the online quizzes used in this investigation was small enough (i.e., each quiz equaled 1% of the final grade) to be considered “nonthreatening.” The online quizzes used herein met many of the most common definitions used to describe formative assessments. For example, the quizzes were of the same exact format as the summative exams (19), they provided early and rapid feedback to the students about how they were doing in the class (1, 19), and they assisted the teacher in modifying the coursework to better meet the students’ needs (17, 19). Therefore, the results of this investigation agree with previous studies that have demonstrated that formative assessments enhance summative exam performance in dental students (15), medical students (2), and a variety of undergraduate major students (5, 6, 14, 21). Furthermore, not only was the effect size pertaining to the benefit of the formative assessments on summative performance similar to that previously reported (15) but the extent of the benefit (~4%) was within the 3% (2) to 6% (14) range of improvements found in the literature.

The results of this investigation also disagree with those studies (3, 11, 17, 18) that found that formative assessments do not enhance summative exam scores. Those latter studies provided two common sets of explanations for their conclusions. The first common explanation was that their students did not sufficiently use the formative assessments because they did not consider them to be worthwhile (17, 18). The second explanation was that student performance on summative assessments suffered because the students attended to the formative assessments in lieu of reading and studying the other course materials (3). It was exactly those types of concerns that shaped the motivation and implementation of the online quizzes that were used in this investigation, which, in turn, contributed to the unique characteristics of this study. To my knowledge, this is the first study that sought to determine if routinely scheduled mandatory (i.e., for credit) online quizzes could be used not just as effective formative assessments but to also enhance student class preparation and participation by encouraging them to read and think critically about the material before class. On the basis of the enhancement in summative exam scores, as well as the evidence from both objective and subjective measures of class preparation and participation discussed above, it is clear the online quizzes were successful in both respects.

**Online quizzes may predict summative exam scores.** The other major outcome of this investigation was the significant positive correlation between individual mean online quiz scores and individual mean summative exam scores in group 3. This finding agrees with those of previous investigations (8, 14, 16) that have also demonstrated that formative assessments show predictive validity for summative exams. These results also reinforce the conclusion that the online quizzes were effective formative assessments because they were similar in nature and they evaluated the same knowledge and summative exams (19). So, in addition to providing valuable feedback to the students about how they are doing, online quizzes may be used by the professor to identify and help those students that are most likely to perform poorly on the summative exams.

**Summary.** Students in an undergraduate Exercise Physiology course that completed formative online quizzes performed better on summative exams than did those students that did not complete the quizzes. The quizzes likely benefited students by motivating them read ahead in the course textbook and lecture notes; thus, they were better prepared for class, they got more out of the class meetings, and, ultimately, they more effectively learned the course material. The online quizzes were also found to have the secondary benefit of being valid predictors of performance on upcoming summative exams.

**REFERENCES**


**How We Teach**

**USE OF ONLINE QUIZZES TO ENHANCE CLASS PREPARATION**

301

Advances in Physiology Education • VOL 32 • DECEMBER 2008

Downloaded from http://advan.physiology.org/ by IP 10.220.33.2 on June 17, 2017
6. Carillo-de-la-Pena MT, Bailles E, Caseras X, Martinez A, Ortet G, Perez J. Formative Assessment and Academic Achievement in Pre-Graduate Students in Health Sciences (online). http://www.springerlink.com/content/t0v831546755w38m/fulltext.html [31 October 2007].