
The use(lessness) of online quizzes for achieving student learning

Harm-Jan Steenhuis* and Brian Grinder

College of Business and Public Administration,
Eastern Washington University,
668 N. Riverpoint Blvd., Suite A,
Spokane, WA 99202, USA
Fax: +1 509 358 2267
E-mail: hsteehuis@mail.ewu.edu
E-mail: bgrinder@mail.ewu.edu
*Corresponding author

Erik Joost de Bruijn

School of Management and Governance,
University of Twente,
P.O. Box 715,
7500 AE, Enschede, The Netherlands
Fax: +31 53 489 2159
E-mail: e.j.debruijn@utwente.nl

Abstract: This paper describes experiences with online quizzes in an operations management course. Online quizzes were introduced to offset larger class sizes. During several quarters, experimentation with online quizzes took place including the number of attempts, the amount of time allowed and the topical coverage in the quizzes. Three research questions are explored: what type of online quiz taking behaviour do students have? Do online quizzes help students improve their grade, that is, Does improvement occur within a course? and, Does the use of online quizzes help a class overall? It was concluded that online quizzes do not significantly improve student learning.

Keywords: online test; online learning; education.

Reference to this paper should be made as follows: Steenhuis, H-J., Grinder, B. and de Bruijn, E.J. (2009) 'The use(lessness) of online quizzes for achieving student learning', *Int. J. Information and Operations Management Education*, Vol. 3, No. 2, pp.119–148.

Biographical notes: Harm-Jan Steenhuis is an Associate Professor of Operations Management at Eastern Washington University and Chair of the Department of Management. He received his MSc in Industrial Engineering and Management and his PhD in International Technology Transfer from the University of Twente, the Netherlands. Currently, he is involved in research on international technology transfer and manufacturing, industry-university technology transfer and instructor-student knowledge transfer.

Brian Grinder is an Associate Professor of Finance at Eastern Washington University. He received his MBA from Fort Hays State University and his PhD from Washington State University. His current research interests include technology in the classroom, finance pedagogy and financial history.

Erik Joost de Bruijn is a Professor of International Management. He received his MSc from the University of Massachusetts and a PhD from the University of Twente. Since 1971, he worked in various industrialisation projects in developing countries. Currently, he teaches International Business Management at the School of School of Management and Governance, University of Twente.

1 Introduction

During the 2005–2006 academic year, a traditional introductory operations management class, which was designed to accommodate a maximum of 60 students, was redesigned so that it could handle more than 150 students. In anticipation of the changing lecture size, online quizzes were introduced in an attempt to move some of the learning from the classroom to the home situation and to add flexibility in learning by allowing students to study and take a quiz when they are prepared. This made the course a hybrid course as described by Flamm et al. (2008).

To evaluate the usefulness of online quizzes three research questions were posed:

- 1 What type of online quiz taking behaviour do students exhibit?
- 2 Do online quizzes help students improve their grades during the course?
- 3 Does the use of online quizzes help a class overall, that is, do courses that use online quizzes enhance student learning more than courses without online quizzes?

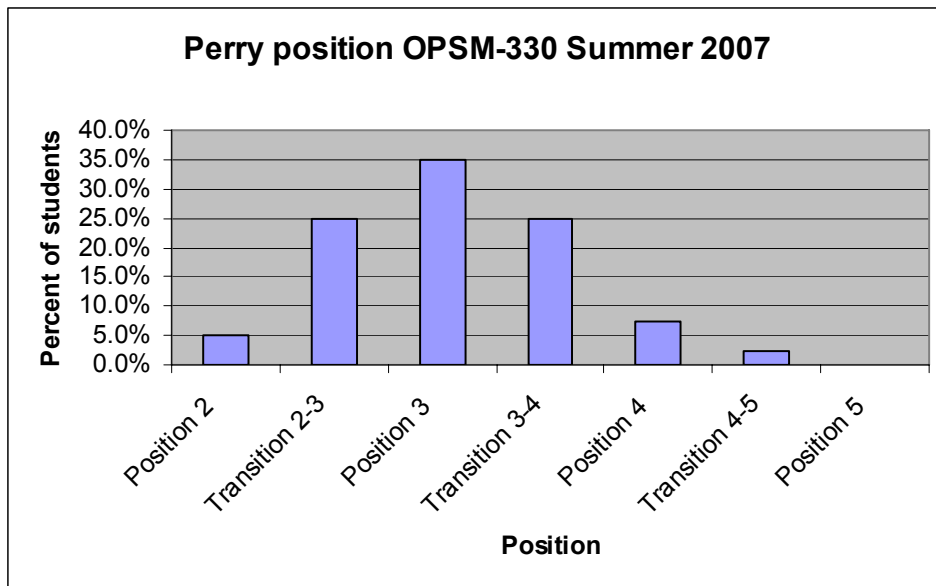
This paper is structured as follows. First, a short introduction is provided about the academic environment in which the course is taught. Section 3 provides a literature overview on the use of modern technologies in the classroom. Section 4 introduces the research setting for the study while sections 5–7 cover, respectively, the first, second and third research question. Finally, in section 8, conclusions are presented.

2 Academic environment

Entwistle and Tait (1990) demonstrated that the academic environment influences how students approach their classroom. Some general background information is therefore provided. The operations management course is a junior level mandatory class for business students. Roughly 300 students a year take this course during the three regular quarters. Additionally, 60–80 students typically take the course during the summer quarter. The course is taught at a regional comprehensive public university that has almost 10,000 students. Around 1,200 students are enrolled in the undergraduate and graduate business programs. Measurements in the operations management course during the spring 2007 quarter, which is considered representative, indicated that roughly 30% of the students could be classified as non-traditional students, approximately 75% of the

students had taken courses at other institutes (many were transfer students from community colleges) and the average GPA at the university when entering the operations management course is a 3.1 (at this university a scale from 0.0 to 4.0 is used). The cognitive development position of students was assessed by using the Learning Environment Preferences instrument developed at the Center for the Study of Intellectual Development¹. This instrument measures students' cognitive development position according to Perry's (1999) model. Perry's model consists of nine positions that fall into four stages. The first stage is dualism in which the student views the world in terms of right and wrong or black and white. In the extreme, this corresponds to position 1 while in position 2 a movement towards the second stage has begun. The second stage is multiplicity. In this stage, students realise that a number of correct answers might exist simultaneously. This corresponds with positions 3 through 4. The next stage is relativism in which a truly relativistic view of the world is adopted. This corresponds with positions 5 and 6. The last stage is the commitment stage corresponding to stages 7–9. In this stage, an increasing commitment to a particular view is made with the corresponding responsibility of that commitment. Assessment in the operations management course took place over several quarters. Results for the summer 2007 quarter are shown in Figure 1. These are similar to measurements in other quarters. Figure 1 shows that most students fall somewhere between positions 2 and 3. The cognitive position of students is one of the factors that might explain class performance as well as provide an overall indication about how students approach the classroom (Wankat and Oreovicz, 1993).

Figure 1 Approximate Perry position according to LEP instrument (see online version for colours)



In the operations management course, instruction follows a learning centred approach similar to what Samuelowicz and Bain (2001) term 'preventing misunderstandings' or 'negotiating misunderstandings'. These approaches require a two-way communication. Additionally, the goal in the course is to achieve deep learning because this is more meaningful than surface learning (Biggs, 1999). The student learning objectives in the course are based on Bloom's taxonomy (Bloom, 1987). They are oriented towards the lowest three learning levels and stated as following:

Knowledge level:

- 1 Know the vocabulary of the operations management discipline.

Comprehension level:

- 2 Describe the functional and supporting roles of operations management in a variety of production and service organisations.
- 3 Explain key operations management concepts.
- 4 Interpret solutions to quantitative data problems relevant to the operations management discipline.

Application level:

- 5 Apply mathematical formulas to quantitative data problems relevant to the operations management discipline.

3 Online learning

The literature on the use of modern information and communication technologies in educational settings can be divided into two distinct categories. Technology can be used for instructional purposes or for testing purposes. The use of an online quiz as a tool to test a student's knowledge and skills is related to the assessment of the student's capabilities. Anderson et al. (2001) identify two different types of assessment of student learning. One type is summative assessment in which information is gathered after the learning should have occurred in order to determine how well the student has learned the material. The other type of assessment is formative assessment. This type is concerned with gathering information about learning as learning is taking place so that timely instructional modifications can be made to improve the quality or amount of learning. The purpose for using online quizzes in the course most aligns with that of formative assessment.

It is often difficult to distinguish summative assessment from formative assessment in the literature because this difference is not always described in the research. By definition, the essential difference is that summative assessment is more a conclusion in the end whereas formative assessment is meant to provide feedback to improve the end result. For this paper, studies where online quizzes are used with the purpose of providing feedback are considered formative in nature.

Students perceive online quizzes as one of the most important pedagogical tools although this perceived helpfulness was not significantly reflected in exam scores (Gurung, 2003). In other words, how students perceive online quizzes and the helpfulness of online quizzes does not necessarily correspond with the actual performance of students in the classroom when measured by exam scores.

It should be noted that, in some instances, online quizzes are used primarily to get students engaged with the material or to influence student attitudes towards the material. Student learning in these instances is a secondary goal (see e.g. Francis and Schreiber, 2008; Nguyen et al., 2006). There are different types of assessment and some have argued that a variety of online assessments should be used in each course (Gaytan and McEwen, 2007). Types of assessment include outcome assessment, process assessment and input assessment (Lucas and Associates, 2000). It goes beyond the scope of this paper to discuss whether student learning outcomes (academic achievement) should be the only determinant of a grade or whether, for example, additional elements such as effort should be part of a grade or not. This is a contentious issue that may also have to do with a wider discussion on grade inflation (see e.g. Gose, 1997; Greenberger et al., 2008; Guskey and Jung, 2006; Holsendolph, 2005; Nature, 2004). The remaining part of this section provides an overview of research on formative quizzes.

Maki and Maki (2001) studied 280 participants in a General Psychology course at Texas Tech University using mastery quizzes. Each quiz contained 15 multiple-choice items and the criterion for passing was 80%. Questions were taken from the test bank that came with the textbook. Students were required to pass two mastery quizzes each week and could take up to two more for additional credit. Students were allowed multiple (possibly unlimited) attempts on each quiz. Every three or four weeks an in-class examination was scheduled covering three or four chapters. This was a paper-and-pencil test that was similar in format to the online quizzes. Questions were also drawn from the test bank, but none of them had appeared on the quizzes. It is not clear how the quiz grades were weighted in relation to the exams, but since the examinations covered the same material, the quizzes can be regarded as formative. Maki and Maki (2001) found that by making the quizzes a requirement, more people took the quizzes (95%) than when they were only voluntary (56%). Maki and Maki (2001) doubt that students used the quizzes for their intended purpose of diagnosing content knowledge. Instead, they found that students printed out the quizzes and used them to prepare for exams, which is not in keeping with their original intent. They also found that there was a significant correlation between performance on master quizzes and on the examinations. However, they note that such correlations are notoriously difficult to interpret. One possible explanation according to them is that students' learning from quizzes is expressed on the examinations but it could also be due to some other 'third' variable (Maki and Maki, 2001). One variable that Maki and Maki (2001) ignore is that the correlation between quiz score and exam score may be due to the innate ability of students. That would mean that the quizzes can be perceived as separate assessments where students do not necessarily learn from the feedback provided on the formative quiz but rather, a good student will perform well on both the quiz and the exam and a less able student will perform poorly on both the quiz and the exam. Then, the mastery quizzes would only serve as an indicator as to what a student's grade will likely be on exams, but the formative element (that would lead to improved scores) would not be present.

Brothen and Wambach (2001) distinguished two types of strategies that students follow when taking quizzes: the quiz-to-learn strategy and the prepare-gather feedback-

restudy strategy. In their course, they used 26 computerised 10-item multiple-choice chapter quizzes (260 possible points) and 4 computerised closed book 50 item unit exams (200 possible points covering six or seven chapters with questions from the chapter quiz bank) which accounted for 75% of the total course points. Chapter quizzes could be taken an unlimited number of times either in the classroom during class or in the campus computer labs. A minimum score of eight was needed for a quiz to count towards a student's grade. The unit exams could only be taken once. The chapter quizzes can be viewed as formative assessment in particular because the unit exams counted more heavily and were based on the same database. Brothen and Wambach's (2001) study included 29 sophomore and junior students in a life span human development course. They found that

- 1 generally, students did better on the chapter quizzes and poorer on unit exams
- 2 students spent an average of 8.67 min on each quiz
- 3 the average student took 132.52 quizzes over the 26 chapters with a mean score of 7.77
- 4 the number of quizzes taken was negatively correlated with mean quiz score
- 5 taking many quizzes did not seem to lead to high unit exam scores.

However, they note that the relation between number of quizzes taken and quiz performance was not uniform (Brothen and Wambach, 2001).

Grimstad and Grabe (2004) focused on the voluntary use of online study resources. Their study included 179 students in an introductory psychology course at a moderate-sized Midwestern state university. Students had access to approximately 100 multiple-choice questions for each chapter for practice. These were selected from the odd-numbered questions from the textbook testbank. There were three exams with 50-item multiple-choice questions covering four textbook chapters. These questions were selected from the even-numbered items in the testbank and some were written by the instructor. The practice questions that could be answered online can therefore be viewed as formative assessment. Grimstad and Grabe (2004) categorised participants who answered at least 50 questions in preparation for an examination as a user of the online study system (12.5% of available questions). For the first, second and third examinations there were, respectively, 31.0%, 38.6% and 32.4% users. These answered a mean of 204, 157 and 285 quiz questions, respectively, for the first, second and third examination. Grimstad and Grabe (2004) found that

- 1 reading ability correlated with exam performance, after accounting for this they also found that
- 2 users scored higher than non-users on examinations with no significant differences between the short delay group users (students immediately moving to the next question when answering incorrectly) vs. the long delay group users (students who delayed considering another question when answering incorrectly)
- 3 the number of questions attempted in preparation for the three examinations and performance on the practice questions 'were either statistically unrelated or positively correlated'.

Grimstad and Grabe (2004) propose that students react differently to voluntary quizzes compared to the mastery quizzes used by Brothen and Wambach (2001). For the voluntary quizzes students may continue to examine items as long as they perceive this to be helpful; for the mastery quizzes, they will do it until they have reached the criterion. Also, for the voluntary quizzes there might be an incentive for students to look up the answer in the book whereas for the mastery quizzes there might not be such an incentive.

Daniel and Broida (2004) looked at the utility of web-based quizzing compared to other types of quizzing. Their study included 125 students enrolled in three sections of Child and Adolescent Development at a public university in New England. The students were assigned to three sections. One section did not use quizzes, one section used in-class quizzes, and one section used web-based quizzes. The in-class quiz group took 16 weekly chapter-based ten (multiple-choice and short answer) question quizzes during the first 15 min of class. Students in the web-based quiz group received the same quizzes. These were available for self-administration on the web 24 hr before class, and students were allowed 15 min for completion. For the first half of the semester both quiz groups received identical quizzes and all three groups completed the same exams. It was not clear how the quizzes were graded compared to the examinations, that is, the weight, but since the examinations covered the same material, the quizzes can be regarded as formative. During the first half of the semester, Daniel and Broida (2004) found that in-class quizzing had a positive impact on the mid-term exam scores, but there was no obvious impact from web-based quizzing on exam performance compared to the no-quiz group. In a search for explanations, Daniel and Broida (2004) found that students in the web-based group reported a number of strategies for cheating such as printing and sharing of quizzes, looking up answers in the book, using an online glossary, etc. Daniel and Broida (2004) changed the system in the second half so that each student's quiz consisted of different questions and the time allowed was reduced from 15 to 7 min. During the second half of the semester, they found that both the in-class group and the web-based group performed better on exams, and there were no significant differences between the in-class group and the web-based group (Daniel and Broida, 2004). Daniel and Broida (2004) conclude that efforts to discourage students in the web-based group from using strategies to optimise their quiz score without mastering the text were successful but note that there was no evidence that online quizzes have benefits over in-class quizzes.

Kibble (2007) used online quizzes in a Medical Physiology course in the first or second year of the program with a cohort of approximately 350 students. The course contained two summative exams after nine and 18 weeks consisting of 80 multiple-choice questions. Two online quizzes were offered during the learning period before each summative examination. The two quizzes offered in each block were presented sequentially, separated by roughly two weeks. The second quiz in the pair included new course material to provide ongoing feedback. Each online quiz consisted of 20–30 multiple-choice questions designed to match the style and difficulty of summative examinations. Different variations were applied to five cohorts from 2004 to 2006, see Table 1.

Table 1 Five different variations for online quizzes

<i>Cohort</i>	<i>Quiz available</i>	<i>Course credit per quiz</i>	<i>Criteria for earning credit</i>	<i>Student participation (%)</i>
1	One week, only one attempt	None		52
2	Unlimited attempts leading up to the summative examination	0.5%	Taking it	87
3	Unlimited attempts during a 1-week window for a given quiz	1.0%	All-or-none basis to students scoring 30% or more on a given quiz	92
4	One week window, 2 attempts	1.0%	Based on the actual quiz score attained from better of two attempts	97
5	One week window, 2 attempts	2.0%	Based on the actual quiz score attained from better of two attempts	98

For the first cohort, Kibble (2007) found that the mean online quiz score did not differ significantly from the mean summative scores. However, individual quiz scores showed a significant positive correlation with corresponding summative examination scores. Students who elected not to use the online quizzes performed significantly worse on the associated summative examination than those who used at least one quiz (Kibble, 2007). For analysis of the other four cohorts Kibble noticed that for cohort 4 and 5, 96% of the students who scored between 95% and 100% on their first quiz attempt, did not take the second quiz, and did not sustain the high level of performance on the summative examination. This suggests an inappropriate use of quizzes to gain credit rather than to provide formative feedback (Kibble, 2007). Therefore, analyses were separated for students who took a quiz once compared to more than once. Kibble (2007) found that:

- 1 The mean quiz score for students using only one attempt was significantly greater than the first effort of students who subsequently repeated a quiz (cohorts 2–5).
- 2a There was no significant difference in summative examination scores for students who took a quiz only once compared with students who took a quiz more than once (cohorts 2, 4).
- 2b The summative examination scores for students who took a quiz only once were significantly lower than students who repeated an online quiz (cohorts 3, 5).
- 3a Individual scores from the first of several quiz attempts showed a significant positive correlation with corresponding summative examination scores (cohorts 2–4).
- 3b Individual scores from the first of several quiz attempts showed no significant correlation with corresponding summative examination scores (cohort 5).

- 4a The mean summative examination score for students who did not take any online quiz was significantly lower than the mean of students taking at least one online quiz (cohorts 2, 3).
- 4b The mean summative examination score for students who did not take any online quiz was not significantly different from the mean of students taking at least one online quiz (cohorts 4, 5) (however, the number of non-participants was very small for these two cohorts).

Overall, it can be concluded from Kibble (2007) that students who participated in the quizzes performed better than students who did not, confirming Grimstad and Grabe's (2004) findings. Kibble (2007) also used a student survey for cohorts 4 and 5 and found that at least 80% of respondents agreed or strongly agreed that they used quizzes in their learning, more than 90% of the respondents agreed that they completed quizzes because credit was associated to them, and over 90% of the respondents agreed that online quizzes could adequately replace offline quizzes. Furthermore, respondents indicated that lecture notes and peer discussion were commonly used to assist in taking quizzes (Kibble, 2007). One conclusion Kibble (2007) draws is that giving credit for the quizzes increased participation but that this may have been reinforced by criteria that made it more difficult to earn full credit, this aligns with Maki and Maki's (2001) findings. Kibble (2007) also found that, contrary to his expectations, large numbers of students took quizzes only once and were able to score close to 100% but were not able to sustain that level on the exams. This, combined with the widespread inappropriate use of online quizzes seems to confirm Brothen and Wambach's (2001) quiz-to-learn strategy as well as Maki and Maki's (2001) and Daniel and Broida's (2004) finding of 'cheating'. For this group of students, Kibble (2007) concludes that any formative value of the quizzes was likely to be lost, defeating his purpose in providing them. Dopper and Sjoer (2004) show similar findings. Students do not necessarily use the formative feedback (by taking action such as consulting fellow students, the teacher or the book), even if this purpose is pointed out to them, and even if they recognise that the feedback is valuable.

Smith's (2007) study included two courses. One course was a lower-division Earth History course while the other was an upper division Environmental Geology course. Smith (2007) used several assessment instruments, including formative online quizzes, and used Bloom's taxonomy for different types of learning that was assessed. In both courses, 10 or 11 weekly quizzes were administered on WebCT; they included 7–12 multiple-choice or short answer questions. The students had 60 minutes or less to complete them and overall, they accounted for 10% or 16% of the grade. Students had the ability to reanswer 1/4–1/3 of incorrect short-answer questions. The quizzes included reflective free-point questions (4–5%) to discuss most significant or least understood concepts for the week. Smith found that online quiz and exam scores correlated strongly with a regression line that had a slope close to one and a *Y*-intercept close to zero, which means that the quiz scores are a nearly exact prediction of exam scores (Smith, 2007). Smith further notes that students who frequently revisited their graded quizzes for feedback and commonly reanswered questions for which they initially did not receive full credit had an overall higher performance compared to those who did not reanswer questions (Smith, 2007). This seems to confirm Kibble's (2007) findings that those students who use the feedback and take quizzes again do better than those students who do not use the feedback. Lastly, Smith (2007) notes that most students who missed more than five classes did not perform well on exams or quizzes.

A study that looked at the effect on exam performance based on when quizzes are accessed is Metz (2008). She looked at weekly graded online quizzes in both an introductory biology course and an upper division biology course (90 and 125 students, respectively) at Montana State University. The main perspective of the study was to see whether cheating occurred. The hypothesis was that the majority of students would access the quiz late in the access period, and that average scores would increase over the access period as early quiz-takers shared information about quiz content with fellow students. For the introductory course, the online quizzes were available for 72 hr. Students could start the quiz at any time during this period and had 20 min to complete the quiz. There were a total of 11 quizzes (in 15 weeks) covering lecture material from the previous week and assigned reading. The lowest quiz score was dropped in the final grade calculation and online quizzes represented 10% of the overall course grade. The set-up for the advanced course was similar except students had only a period of 24 hr to take a quiz, 12 quizzes were administered, only the top-nine were used for grade calculations, and they accounted for 12% of the course grade. The quizzes contained 10 questions that were pulled from a quiz bank with 15–30 questions depending upon the course and week. Unique quizzes were generated for each student and for multiple-choice items, the order of the answer choices was scrambled each time a quiz was generated. Questions were designed by the instructor to cover varying levels of Bloom's taxonomy educational objectives with three of the 10 questions requiring higher order thinking skills. Metz (2008) found that students taking quizzes later in the access period were more likely to achieve lower scores than students taking quizzes early. Therefore, she concludes that widespread cheating on the online quizzes does not occur. It should be noted that this conclusion contradicts Daniel and Broida's (2004) conclusion about cheating (see earlier). Metz (2008) also found that quizzes correlated positively with overall course achievement with a nearly 1:1 correspondence. A comparison of the course with online quizzes and a similar course (same instructor same types of quizzes) with in-class quizzes revealed no differences between the quizzes administered during lecture and those administered online. Metz (2008) further found that the largest block of students took the quiz in the last hour before it was due, whether the quiz access closing time was 3 p.m. or 8 a.m. This finding is similar to a study by Taraban et al. (1999) who found that students typically only engage with the material a day or so before an exam, view this cramming as an ideal strategy, and spend very little time outside the classroom on material (30 min a day student self-reported). Scores for quizzes taken from midnight to 8 a.m. were significantly lower than for quizzes taken during other times of the day, while there was no difference in average scores for quizzes taken during daytime (8 a.m.–4 p.m.) or evening (4 p.m.–midnight). Metz (2008) provides two potential explanations. Firstly, students who are night-owls by nature may tend to do worse in courses overall because they have difficulty engaging in courses taught during daytime hours. Secondly, students who spend many hours working, have time consuming non-academic responsibilities (e.g. families) or are taking very large course-loads may be forced into completing online assignments late at night. For these students, fatigue may play a role in their quiz performance. Metz (2008) further notes that a survey on student perceptions revealed that the time constraints placed on completion of the quiz were considered too stressful or unfair.

4 The setting

The course redesign with online quizzes was first implemented in the spring 2006 quarter. It included a number of technological enhancements including the use of Perception software to administer the online quizzes and a clicker system to enhance classroom discussions. Online quizzes with multiple attempts were originally introduced as a feedback mechanism where students would study, take a quiz, receive feedback, study again, retake the quiz and hopefully improve their scores. This is the prepare-gather feedback-rewrite strategy as defined by Brothen and Wambach (2001). The operations management course can be classified as a web-enhanced course according to Driver's (2000) classification schema. Regular in-class paper and pencil exams were used in addition to the online quizzes. One exam was given at the middle of the quarter and covered the first half of the materials; the second exam was given at the end of the quarter and covered the second half of the materials. In accordance with the course objectives, the exams contained questions on each of the three Bloom's taxonomy levels addressed in the course.

To discourage simply taking quizzes multiple times and hoping for a better result and to encourage students to study between quizzes, the last online quiz that was taken was used for grading purposes. Quizzes contained 10 questions from test banks that came with the textbook. Roughly 30–50 questions were available per chapter and typically 1 or 2 chapters were covered on the online quiz. Quizzes were given approximately one per week a week and were available to students for 5–7 days depending on the quarter. Table 2 provides some general information on the course, the number of students and the quiz characteristics.

Table 2 Characteristics of online quizzes

<i>Quarter</i>	<i>Students</i>	<i>No. of quizzes</i>	<i>Quest per quiz</i>	<i>Minutes per quiz</i>	<i>Course credit per quiz (%)</i>	<i>Quiz window</i>	<i># attempts allowed</i>
Spring 2006	63	9	10	20 for math 10 for theory	3.3	7 days	After the 1st week 5
Summer 2006	42	7	10	16 mix	2.9	5 days	5
Fall 2006	95	9	10	18 theory	0.9	6 days	5
Spring 2007	135	6	10	18 theory	1.4	6 days	5
Summer 2007	54	5	10	18 theory	0.5	6 days	5
Fall 2007	112	Not used anymore					
Spring 2008	124	Not used anymore					
Summer 2008	70	Not used anymore					

5 Quiz taking behaviour

In this section, the first research question will be discussed: What type of online quiz taking behaviour do students exhibit? This discussion will focus on when students take quizzes, the number of attempts they take, how long they take to do a quiz, and the time between attempts.

5.1 First quarter experiences

During the spring 2006 quarter, students were initially allowed to take unlimited attempts on quizzes but during the very first quiz, some students took over 30 attempts essentially depleting the entire test bank. Therefore, in the second week a limit of five attempts was introduced. Also, during spring 2006, quizzes alternated in their orientation between theory and math. These weekly online quizzes became available to students for one week on Friday at noon until the next Friday at 8am. Some exceptions occurred due to holidays.

5.1.1 Number of attempts

Table 3 shows the percentage of students by quiz and how many attempts they took.

Table 3 Percentage of students and number of attempts, spring 2006

	<i>0 attempts</i>	<i>1 attempt</i>	<i>2 attempts</i>	<i>3 attempts</i>	<i>4 attempts</i>	<i>5 attempts</i>	<i>6 or more attempts</i>
Quiz 1	4.8	12.7	6.3	14.3	11.1	7.9	42.9
Quiz 2	11.1	28.6	27.0	12.7	7.9	12.7	
Quiz 3	6.3	23.8	15.9	31.7	12.7	9.5	
Quiz 4	4.8	15.9	36.5	14.3	14.3	14.3	
Quiz 5	3.2	11.1	25.4	14.3	23.8	22.2	
Quiz 6	7.9	22.2	25.4	28.6	6.3	9.5	
Quiz 7	7.9	36.5	33.3	12.7	6.3	3.2	
Quiz 8	12.7	22.2	28.6	20.6	14.3	1.6	
Quiz 9	4.8	34.9	31.7	14.3	11.1	1.6	1.6

5.1.2 When attempts are taken

Tables 4 and 5 show when students took their first and last attempt on the first two quizzes and the last two quizzes of the quarter, these are considered representative. For both these tables, for ease of analysis, the time slots are divided into 12 hr blocks, that is, from midnight until noon and from noon until midnight.

Table 4 Timing of the first attempt for four different quizzes (in % of students)

<i>Time period</i>	<i>Quiz 1</i>	<i>Quiz 2</i>	<i>Quiz 8</i>	<i>Quiz 9</i>
Day 1: noon until midnight	3.1	3.5	3.6	3.3
Day 2: midnight until noon	1.6	1.8	3.6	0.0
Day 2: noon until midnight	3.1	3.5	0.0	5.0
Day 3: midnight until noon	0.0	0.0	0.0	0.0
Day 3: noon until midnight	7.8	0.0	3.6	16.7
Day 4: midnight until noon	1.6	3.5	0.0	3.3
Day 4: noon until midnight	9.4	0.0	12.7	13.3
Day 5: midnight until noon	1.6	1.8	3.6	8.3
Day 5: noon until midnight	9.4	3.5	20.0	5.0
Day 6: midnight until noon	3.1	1.8	1.8	1.7
Day 6: noon until midnight	14.1	22.8	47.3	23.3
Day 7: midnight until noon	7.8	8.8	3.6	6.7
Day 7: noon until midnight	32.8	42.1		11.7
Day 8: midnight until 8 a.m.	4.7	7.0		1.7

Table 5 Timing of the last attempt for four different quizzes (in % of students)

<i>Time period</i>	<i>Quiz 1</i>	<i>Quiz 2</i>	<i>Quiz 8</i>	<i>Quiz 9</i>
Day 1: noon until midnight	3.13	1.75	1.82	1.67
Day 2: midnight until noon	1.56			
Day 2: noon until midnight	3.13	3.51		1.67
Day 3: midnight until noon			1.82	
Day 3: noon until midnight	1.56	1.75		13.33
Day 4: midnight until noon	1.56		10.91	
Day 4: noon until midnight	3.13			18.33
Day 5: midnight until noon			20.00	5.00
Day 5: noon until midnight	10.94	1.75	3.64	
Day 6: midnight until noon	3.13	1.75	56.36	3.33
Day 6: noon until midnight	10.94	21.05	5.45	5.00
Day 7: midnight until noon	6.25	7.02		1.67
Day 7: noon until midnight	45.31	50.88		23.33
Day 8: midnight until noon	9.38	10.53		10.00
Day 8: noon until midnight				16.67

5.1.3 Time spent per attempt

Table 6 provides insight on the average amount of time used for the same four quizzes.

It should be noted that the time for the 5th attempt is easily influenced by the behaviour of one or a few students since only a few students had that many attempts. Also, it shows that for theory quizzes (quiz 1) the amount of time needed was somewhere between 5 and 6 min whereas for math oriented quizzes students took roughly between 14 and 16 min.

Table 6 Average time spent on quiz in minutes

	<i>1st attempt</i>	<i>2nd attempts</i>	<i>3rd attempts</i>	<i>4th attempts</i>	<i>5th attempts</i>
Quiz 1	5.57	5.45	5.53	5.71	5.82
Quiz 2	15.36	16.08	14.63	15.46	13.63
Quiz 8	15.09	16.46	14.91	13.11	4.00
Quiz 9	15.10	15.62	15.78	12.67	20.00

5.1.4 Time between attempts

Table 7 shows the median time between attempts in the course for four of the quizzes.

Table 7 Median time spent between attempts in minutes

	<i>1st–2nd attempt</i>	<i>2nd–3rd attempt</i>	<i>3rd–4th attempt</i>	<i>4th–5th attempt</i>
Quiz 1	2.5	2.0	1.0	2.5
Quiz 2	39.0	7.5	5.0	4.0
Quiz 8	32.0	13.0	7.0	
Quiz 9	25.0	8.0	6.0	648.5 (<i>n</i> = 1)

5.1.5 Discussion

Tables 3–7 show a variety of things. The following are highlights from these tables:

- On average 7% of the students did not take an online quiz. This is similar to Kibble's (2007) findings.
- After limiting the number of attempts to five, roughly 50–60% of the students only took one or two attempts. This is similar to Kibble's (2007) findings.
- More than a third of the students take their first attempt within the last day that the quiz is available. Although, the time period in this study for taking quizzes is longer than Metz's (2008) study and not tracked per hour, this is similar to Metz's (2008) finding that a large block of students take the quizzes towards the end of the quizzing period. This also confirms Taraban et al. (1999) who conclude that students typically only engage with the material a day or so before an exam.
- Students spent roughly 5.5 minutes on theory quizzes, and the amount of time taken does not vary significantly by attempt.
- Students spent roughly 15 minutes on math quizzes, and there is no clear relationship between the time taken on a quiz and the number of attempts.
- The median time between attempts is usually small indicating that students are not taking the time to study their book again before they take an additional attempt. This

indicates that students tended to use the quiz-to-learn strategy (Brothen and Wambach, 2001; Dopfer and Sjoer, 2004; Kibble, 2007) and is similar to Maki and Maki's (2001) finding that the quizzes were not used as intended.

5.2 Second quarter experiences

Starting in summer 2006, quizzes were changed to a mix of theory and math questions. Based on evidence from spring 2006 (see Tables 3–7), it was determined that students did not use the online quizzes as they were envisioned by the instructor when they were incorporated in the course design (Brothen and Wambach's (2001) prepare-gather feedback-restudy strategy). The quiz taking behaviour patterns indicated that students did not use the feedback from the online quiz performance to study the material again and subsequently retake the quiz with better preparation. It was fairly common for students to retake a quiz almost immediately after their most recent attempt (see Table 7) indicating a strong inclination towards the quiz-to-learn strategy (Brothen and Wambach, 2001; Kibble, 2007).

To motivate students to use the prepare-gather feedback-restudy strategy while also discouraging the quiz-to-learn strategy, a different grading schema was devised. The total quiz score was based on a formula:

$$\text{Quiz score} = \text{Effectiveness} \times \text{Efficiency} \times \text{Attempts}$$

In this formula, the variables are determined as follows:

- Effectiveness = number of correct answers (out of 10)
- Efficiency is based on the total time needed to take the quiz
 - Up to 12 min = 1.0
 - 12–14 min = 0.9
 - 14–16 min = 0.8
- Attempts
 - 1 attempt = 1.0
 - 2 attempts = 0.95
 - 3 attempts = 0.9
 - 4 attempts = 0.85
 - 5 attempts = 0.8

The idea behind the efficiency multiplier was to encourage students to prepare beforehand by studying for the quiz rather than looking up answers during the quiz. From Table 6, it follows that in spring the theory quizzes took about 5.5 min and the math quizzes about 15 min leading to a total of 20.5 min for 20 questions. Adjusting down to 10 questions leads to roughly 11 min per quiz. This was rounded up to 12 to allow for extra flexibility. Some additional time was allowed, that is, going up to 16 min, but the intent of the time limit was to discourage students from trying to look up answers in the book while taking the quiz. Daniel and Broida (2004) also noted this type of behaviour.

The attempt multiplier was designed to discourage students from simply taking a quiz again and gambling/hoping for a better score on the subsequent attempt and to encourage students to study before taking another attempt. The idea was that by including this

multiplier, a cost would be imposed on students who did not study before taking the quiz again.

During summer 2006, quizzes were available from Tuesday at 8pm until Sunday at midnight, a little more than five days. It was changed to Tuesday at 8pm to coincide with the ending of the lecture and it was left open until Sunday at midnight to allow students to finish it on the weekend.

5.2.1 Number of attempts

Table 8 shows the number of attempts during summer 2006.

Table 8 Number of attempts, Summer 06

	<i>0 attempts (%)</i>	<i>1 attempt (%)</i>	<i>2 attempts (%)</i>	<i>3 attempts (%)</i>	<i>4 attempts (%)</i>	<i>5 attempts (%)</i>
Quiz 1	0	52.4	23.8	19.0	0	4.8
Quiz 2	16.7	42.9	23.8	7.1	9.5	
Quiz 3	0	42.9	33.3	14.3	9.5	
Quiz 4	2.4	54.8	14.3	19.0	9.5	
Quiz 5	2.4	54.8	28.6	11.9	2.4	
Quiz 6	2.4	40.5	21.4	28.6	2.4	4.8
Quiz 7	4.8	52.4	23.8	7.1	7.1	4.8

5.2.2 When attempts are taken

Tables 9 and 10 show when students took their first and last attempt. Similar to Section 5.1.2, the time slots are divided in 12-hr periods.

Table 9 Timing of the first attempt for 7 quizzes (in % of students), summer '06

<i>Time period</i>	<i>Quiz 1</i>	<i>Quiz 2</i>	<i>Quiz 3</i>	<i>Quiz 4</i>	<i>Quiz 5</i>	<i>Quiz 6</i>	<i>Quiz 7</i>
Day 1: 8a-noon	–	–	–	–	–	–	–
Day 1: noon–midnight					2.4	14.6	15
Day 2: midnight–noon			2.3				
Day 2: noon–midnight	2.2	10.8	5.4				7.5
Day 3: midnight–noon						2.4	2.5
Day 3: noon–midnight	8.9	5.4	5.4	4.9	2.4	7.3	7.5
Day 4: midnight–noon		2.7			4.8		
Day 4: noon–midnight	4.4			14.6	14.3	4.9	
Day 5: midnight–noon	2.2				2.4	2.4	5.0
Day 5: noon–midnight	6.7	8.1	10.8	2.4	11.9	2.4	2.5
Day 6: midnight–noon	4.4	8.1	5.4	4.9	2.4	4.9	
Day 6: noon–midnight	11.1	64.9	74.4	73.2	59.5	61.0	60
^a Day 7: midnight–noon	24.4						
^a Day 7: noon - midnight	35.6						

^aThis portion of the quizzing cycle only applied to quiz 1 in Summer 06.

Table 10 Timing of the last attempt for 7 quizzes (in % of students), summer '06

<i>Time period</i>	<i>Quiz 1</i>	<i>Quiz 2</i>	<i>Quiz 3</i>	<i>Quiz 4</i>	<i>Quiz 5</i>	<i>Quiz 6</i>	<i>Quiz 7</i>
Day 1: 8a–noon							
Day 1: noon–midnight					2.4	14.6	15.0
Day 2: midnight–noon			2.3				
Day 2: noon–midnight	2.2	8.1	2.3				5.0
Day 3: midnight–noon						2.4	2.5
Day 3: noon–midnight	8.9	2.7	6.9	2.4	2.4	7.3	7.5
Day 4: midnight–noon				2.4			
Day 4: noon–midnight	4.4			12.2	16.7	4.9	
Day 5: midnight–noon	2.2				2.4		5.0
Day 5: noon–midnight	6.7	5.4	4.7	2.4	14.3	4.9	2.5
Day 6: midnight–noon	4.4	2.7	7.0	4.9	2.4	4.9	
Day 6: noon–midnight	8.9	81.1	76.7	75.6	59.5	61.0	62.5
^a Day 7: midnight–noon	22.2						
^a Day 7: noon – midnight	40.0						

^aThis portion of the quizzing cycle only applied to quiz 1 in Summer 06.

5.2.3 Time spent per attempt

Table 11 provides insight into the average amount of time used by attempt.

Table 11 Average time spent on quiz in minutes, summer 2006

	<i>1st attempt</i>	<i>2nd attempt</i>	<i>3rd attempt</i>	<i>4th attempt</i>	<i>5th attempt</i>
Quiz 1	9.68	8.95	7.7	9	10.5
Quiz 2	11.95	11.11	10	9.5	
Quiz 3	11.88	10.64	10.1	10.75	
Quiz 4	12.05	12.33	10.55	12	
Quiz 5	9.82	10.47	10	15	
Quiz 6	12.71	12.57	11.07	13.67	10.5
Quiz 7	12.41	12.56	11.5	13.4	11.5

5.2.4 Time between attempts

Table 12 illustrates the time between attempts. This is an indication of how quickly students take an additional attempt.

Table 12 Median time spent between attempts in minutes, summer 2006

	<i>1st–2nd attempt</i>	<i>2nd–3rd attempt</i>	<i>3rd–4th attempt</i>	<i>4th–5th attempt</i>
Quiz 1	267.5	350.5	336	140.5
Quiz 2	5	10	1.5	
Quiz 3	7	18	1	
Quiz 4	11.5	10	3.5	
Quiz 5	5	16	2	
Quiz 6	9	3	2	1.5
Quiz 7	6	1	1	5

5.2.5 Discussion

Tables 8–12 provide data on a number of different issues. Although a number of different analyses can be performed, the following can be considered important highlights:

- On average 4% of the students did not take an online quiz. This percentage is similar to Kibble's (2007) findings. It is an improvement compared to spring 2006 even though the total weight of online quizzes was reduced from spring 2006 (30%) to summer 2006 (20%). This last aspect contradicts Kibble's (2007) findings where essentially more weight led to higher participation rates.
- Roughly 50% of the students only take one attempt. Roughly, 70% or more of the students take only one or two attempts. Compared to spring 2006 these percentages went up.
- Roughly 60% or more of the students take their first attempt within the last day that the quiz is available. Compared to spring 2006, this percentage went up.
- Students spent roughly 11 min on a quiz. There is no clear relationship between the time taken and the number of attempts. Compared to spring 2006, this is what was expected by mixing the theory and math questions. In some instances, averages were higher than the 12 min allowed for receiving the full efficiency score.
- The median time between attempts is still mostly small and not significantly different from spring 2006. It is not clear why the numbers for the first quiz are different than for the other quizzes.

5.3 Fifth (last) quarter experiences

In fall 2006, based on anecdotal evidence (some students visiting the instructor's office hours with questions), the quizzes were changed to theory quizzes only. By that time it appeared that students made print-outs of the quizzes (see Daniel and Broida (2004) who considered this cheating), used them as study guides for exams (see Maki and Maki

(2001) with similar experiences), and knew what the right answers were but often did not know why and did not take steps to find out why. This confirms Dopper and Sjoer's (2004) conclusion that students do not necessarily use feedback when taking multiple attempts. One interpretation of the mismatch between knowing the right answer but not knowing why it is the right answer is that the online quizzes, which are based on testbank questions, may have inadequately assessed the appropriate learning level. Or, the students are oriented towards memorisation (what and right vs. wrong) rather than application (why). These anecdotal findings also confirm some of the issues with multiple-choice tests as pointed out by Wijekumar et al. (2006).

To deal with this issue, the weight of the online quizzes was reduced. The purpose for using the quizzes was changed from serving as a study tool based on feedback to an engagement tool that would encourage students to interact with the material at some level before class. The efficiency factor was adjusted 2 min upwards, that is, students had two extra minutes allowing 14 min for a 100% score and 18 min for an 80% score. This extra time was added because the quizzes were no longer viewed as a formative assessment tool but as a tool to encourage students to look at the book. Online quizzes were used in this fashion for three quarters in a row with essentially the same set-up. After that, it was decided that it was not useful to have online quizzes based upon looking up information and the online quizzes were removed from the course. Tables 13–17 provide data for the summer 2007 quarter, that is, the last quarter that the online quizzes were used. Quizzes were available from Monday at 6pm (after the lecture) until Sunday at midnight.

5.3.1 Number of attempts

Table 13 shows the number of attempts that students took for quizzes in summer 2007.

Table 13 Percentage of students and number of attempts, summer 2007

	<i>0 attempts</i>	<i>1 attempt</i>	<i>2 attempts</i>	<i>3 attempts</i>	<i>4 attempts</i>	<i>5 attempts</i>
Quiz 1	9.3	57.4	25.9	3.7	3.7	
Quiz 2	18.5	66.7	9.3	1.9	3.7	
Quiz 3	5.6	70.4	22.2	1.9		
Quiz 4	16.7	66.7	13.0	3.7		
Quiz 5	11.1	64.8	14.8	9.3		

5.3.2 When attempts are taken

Table 14 illustrates when students took their first attempt while Table 15 illustrates when students took their last attempt on a quiz.

Table 14 Timing of the first attempt for quizzes, summer 2007

<i>Time period</i>	<i>Quiz 1</i>	<i>Quiz 2</i>	<i>Quiz 3</i>	<i>Quiz 4</i>	<i>Quiz 5</i>
Day 1: 6 pm–midnight		2.3		2.2	
Day 2: midnight–noon	2.0				
Day 2: noon–midnight		2.3	5.9	4.3	
Day 3: midnight–noon	6.1	2.3	3.9		4.2
Day 3: noon–midnight	4.1	2.3	2.0	4.3	8.3
Day 4: midnight–noon			3.9		
Day 4: noon–midnight		20.5	15.7	21.7	18.8
Day 5: midnight–noon	2.0	2.3	2.0	2.2	2.1
Day 5: noon–midnight	26.5	2.3	2.0	4.3	2.1
Day 6: midnight–noon	8.2	4.5	5.9	2.2	6.3
Day 6: noon–midnight	8.2	18.2	15.7	17.4	12.5
Day 7: midnight–noon		9.1	9.8	2.2	2.1
Day 7: noon–midnight	42.9	34.1	33.3	39.1	43.8

Table 15 Timing of the last attempt for quizzes, summer 2007

<i>Time period</i>	<i>Quiz 1</i>	<i>Quiz 2</i>	<i>Quiz 3</i>	<i>Quiz 4</i>	<i>Quiz 5</i>
Day 1: 6 pm–midnight				2.2	
Day 2: midnight–noon	2.0				
Day 2: noon–midnight		2.3	5.9	4.3	
Day 3: midnight–noon	4.1	2.3	3.9		4.2
Day 3: noon–midnight	4.1	2.3	2.0	4.3	8.3
Day 4: midnight–noon			3.9		
Day 4: noon–midnight		20.5	15.7	21.7	18.8
Day 5: midnight–noon	2.0	2.3	2.0	2.2	2.1
Day 5: noon–midnight	22.4	2.3	2.0	4.3	2.1
Day 6: midnight–noon	8.2	6.8	5.9	2.2	6.3
Day 6: noon–midnight	10.2	18.2	15.7	17.4	10.4
Day 7: midnight–noon	2.0	9.1	9.8	2.2	2.1
Day 7: noon–midnight	44.9	34.1	33.3	39.1	45.8

5.3.3 Time spent per attempt

Table 16 provides insight on the average amount of time used for the quizzes during summer 2007.

Table 16 Average time spent on quiz in minutes, summer 2007

	<i>1st attempt</i>	<i>2nd attempt</i>	<i>3rd attempt</i>	<i>4th attempt</i>	<i>5th attempt</i>
Quiz 1	9.5	8.41	8.5	14	
Quiz 2	8.88	11.11	6.67	6	
Quiz 3	8.5	7.58	11		
Quiz 4	8.36	6.6	5.5		
Quiz 5	10.13	8.46	6		

5.3.4 Time spent between attempts

Table 17 shows the median time between attempts and therefore provides an indication for how much time is available for studying in between attempts.

Table 17 Median time spent between attempts in minutes, summer 2007

	<i>1st–2nd attempt</i>	<i>2nd–3rd attempt</i>	<i>3rd–4th attempt</i>	<i>4th–5th attempt</i>
Quiz 1	2.0	1.0		
Quiz 2	4.0	14.0	3.0	
Quiz 3	9.0	6.0		
Quiz 4	3.0	1.5		
Quiz 5	3.0	1.0		

5.3.5 Discussion

Tables 13–17 provide information for the summer 2007 quarter. The following can be considered highlights for this quarter.

- On average 12.2% of the students did not take an online quiz. This percentage is much higher than in the previous quarters. The total weight of online quizzes reduced from spring 2006 (30%) to summer 2006 (20%) to summer 2007 (2.7%). Kibble (2007) found that if online quizzes count for 2% of the grade, 13% of the students did not participate. The numbers presented above are similar.
- Roughly 65% of the students only take one attempt. Roughly, 80% or more of the students take only one or two attempts. Compared to summer 2006 these percentages went up.
- Roughly 40% of the students take their first attempt within the last day that the quiz is available. Compared to summer 2006, this percentage went down.
- Students spent roughly 8.5 min on a quiz. There is no clear relationship between the time taken and the number of attempts. Compared to summer 2006, this went down. This is not unexpected since these quizzes did not contain math anymore. However, compared to the spring 2006 theory quizzes this number went up. This might have been because more time was available allowing students to look up answers in the book.
- The median time between attempts is still mostly small.

5.4 Conclusions

The data from the previous sections illustrates that the online quizzes did not work as intended. The intended goal was for students to follow the prepare-gather feedback-restudy strategy but instead it appears that students followed the quiz-to-learn strategy. Many students took only one attempt at a quiz. This demonstrates that feedback is not used to restudy and improve performance. Many students also took online quizzes very late in the available time period. This simply did not provide them with an opportunity to restudy. Even if multiple attempts were taken, they typically occurred very shortly after an earlier attempt. This, demonstrates that time was not taken to restudy the material. Furthermore, students seemed to use their books during the online quizzes. These combined findings confirm in a setting that has not been studied before (a business school setting) the observations from Brothen and Wambach (2001), Daniel and Broida (2004), Dopfer and Sjoer (2004), Kibble (2007), Maki and Maki (2001), Metz (2008) and Taraban et al. (1999).

To motivate students to follow the prepare-gather feedback-restudy strategy a change in the grading mechanism was introduced. A comparison of summer 2006 with spring 2006 showed that more students took only one or two attempts and more students took their first attempt during the last day. It can therefore be concluded that this change did not have the desired effect on student behaviour.

Participation rates fluctuated during the different quarters. Kibble (2007) found that increasing the online quiz weight overall from 0% to 8% led to an increase in participation rates from 52% to 98%. The findings above show that participation was 87.8%, 96% and 93% for overall weight of the online quizzes of, respectively, 2.7%, 20% and 30%. This shows that increasing the weight of online quizzes does not necessarily lead to higher participation rates. Instead, there may be a tipping point. Based upon the numbers found and Kibble's (2007) study it seems that a desired strategy would be to have a 4–8% weight which can be expected to lead to a 95% or higher participation rate.

6 Performance over time

This section discusses the second research question: Do online quizzes help students improve their grade? The focus for this research question is whether improvement takes place within a course. This discussion will focus on the score on the first attempt, improvements between attempts, the score on the last attempt, and whether online quiz scores are related to exam scores and/or the final grade.

6.1 Findings

The first analysis examines improvements when students take multiple attempts. Table 18 shows the average improvement from one attempt to the next across the three quarters. This table illustrates that, on average, improvement is inconsistent. In some cases, there is improvement from one attempt to the next while in other cases there is actually a worsening of results.

Table 18 Average improvement on quiz score in percentages across three quarters

	<i>1st to 2nd attempt (%)</i>	<i>2nd to 3rd attempt (%)</i>	<i>3rd to 4th attempt (%)</i>	<i>4th to 5th attempt (%)</i>
Spring 2006: Quiz 1	-4.1	6.9	5.8	1.3
Spring 2006: Quiz 2	10.0	-5.0	6.8	8.3
Spring 2006: Quiz 8	4.6	2.2	3.9	4.4
Spring 2006: Quiz 9	-5.4	6.0	-14.4	-5.0
Summer 2006: Quiz 1	-4.4	5.7	-48.7	45.0
Summer 2006: Quiz 2	-3.1	-7.5	22.2	
Summer 2006: Quiz 3	2.9	2.9	7.5	
Summer 2006: Quiz 4	-2.4	-4.2	10.0	
Summer 2006: Quiz 5	1.2	10.0	-3.3	
Summer 2006: Quiz 6	-6.6	0.7	-5.3	35.0
Summer 2006: Quiz 7	-4.9	-8.1	3.7%	-5.0
Summer 2007: Quiz 1	5.9	-5.0	-25.0%	
Summer 2007: Quiz 2	-7.8	-6.3	20.0%	
Summer 2007: Quiz 3	1.6	0.8		
Summer 2007: Quiz 4	-4.6	10.0		
Summer 2007: Quiz 5	-4.3	8.5		

A second analysis distinguishes repeaters from non-repeaters. In Table 19, a distinction is made between these two types of students and their performance is compared. Table 19 shows that non-repeaters typically perform much better on their first online quiz attempt than the repeaters. Also, Table 19 shows that when looking at the online quiz score students eventually get, there is not much difference between non-repeaters and repeaters. In some instances, the repeaters score higher, see for example quizzes 8 and 9 in spring 2006. In other cases, the non-repeaters score higher, see for example quizzes 1 and 3 in summer 2007.

Another indicator of whether online quizzes 'work' is to analyse how the performance of students on online quizzes is related to the performance of students in the course overall, that is, course grade. Figure 2 shows this relationship for the summer 2006 quarter as an example. A similar analysis was conducted for the spring 2006 and summer 2007 quarters. The correlations for spring 2006, summer 2006 and summer 2007 were respectively: 0.7467, 0.4767 and 0.2767.

Table 19 Improvement from first to last attempt, all three quarters

	<i>First online quiz score non-repeaters</i>	<i>Percentage of repeaters</i>	<i>First online quiz score repeaters</i>	<i>Last online quiz score repeaters</i>
Spring 2006: quiz 1	8.38	86.7	5.33	8.19
Spring 2006: quiz 2	7.89	67.9	4.68	7.92
Spring 2006: quiz 8	7.14	74.5	4.93	7.24
Spring 2006: quiz 9	7.36	63.3	4.92	7.45
Summer 2006: quiz 1	7.41	47.6	5.20	6.75
Summer 2006: quiz 2	7.33	48.6	4.24	6.41
Summer 2006: quiz 6	7.18	58.5	5.38	7.04
Summer 2006: quiz 7	7.14	45.0	4.78	6.33
Summer 2007: quiz 1	8.55	36.7	5.50	8.33
Summer 2007: quiz 2	8.69	18.2	7.16	9.13
Summer 2007: quiz 3	8.87	25.5	4.54	8.00
Summer 2007: quiz 4	8.66	17.4	7.50	8.50
Summer 2007: quiz 5	8.17	27.1	6.00	8.38

Figure 2 Relationship between overall performance on online quizzes and the course grade (see online version for colours)

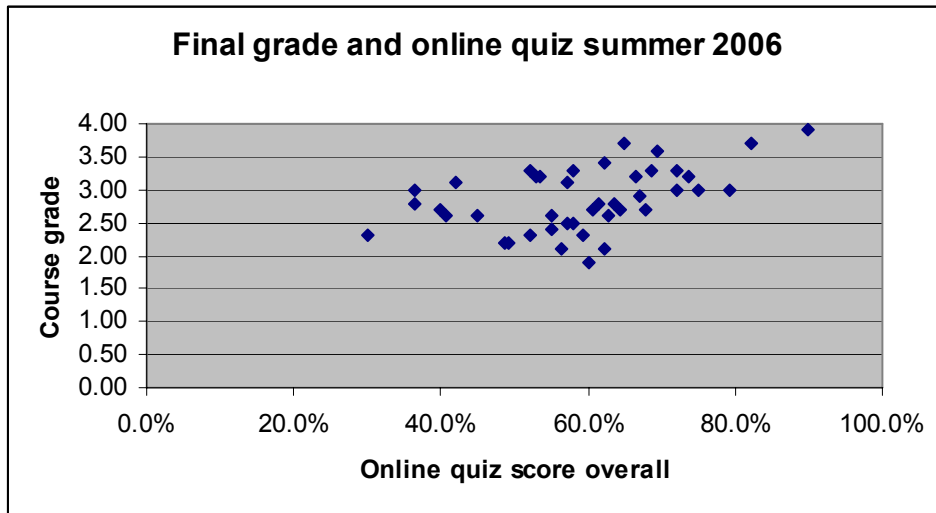


Figure 2 shows that there is a linear relationship between online quiz performance and the course grade. However, this is not unexpected since the online quizzes are part of the grade. Furthermore, the correlation fell dramatically in the subsequent quarters. This is also not unexpected since the online quizzes counted for roughly 30%, 20% and 2.5% of the final grade, respectively, in spring 2006, summer 2006 and summer 2007.

The relationship between online quiz performance and course grade is also not a good indicator of whether the use of online quizzes leads to improved student learning. It is possible that both online quizzes as well as the course grade reflect a student's ability regardless of whether online quizzes help students improve learning.

Lastly, a better indicator of whether online quizzes lead to improved student learning is the comparison of online quiz results with exam results that cover the same materials. The data for the three quarters is provided in Table 20.

Table 20 Online quiz performance and exam performance, all three quarters

<i>Number of students</i>	<i>Spring 2006</i>	<i>Summer 2006</i>	<i>Summer 2007</i>
<i>First half of the quarter</i>			
<i>Students that did not repeat any quiz</i>			
Average score first set of quizzes (in %)	n.a.	68.4	78.7
Average score mid-term exam (in %)	n.a.	71.6	72.8
Average score final grade (max is 4.0)	n.a.	2.78	2.78
<i>Students that repeated some of the quizzes</i>			
Average score first set of quizzes (in %)	72.1	68.5	74.8
Average score mid-term exam (in %)	68.6	73.0	68.2
Average score final grade (max is 4.0)	2.36	2.87	2.56
<i>Students that repeated all of the quizzes</i>			
Average score first set of quizzes (in %)	77.8	65.0	86.7
Average score mid-term exam (in %)	68.4	58.5	66.7
Average score final grade (max is 4.0)	2.55	2.75	2.12
<i>Second half of the quarter</i>			
<i>Students that did not repeat any quiz</i>			
Average score second set of quizzes (in %)	48.6	63.1	77.4
Average score final exam (in %)	43.6	64.0	64.4
Average score final grade (max is 4.0)	2.18	2.85	2.66
<i>Students that repeated some of the quizzes</i>			
Average score second set of quizzes (in %)	71.1	70.3	71.0
Average score final exam (in %)	48.6	68.8	65.4
Average score final grade (max is 4.0)	2.49	2.93	2.71
<i>Students that repeated all of the quizzes</i>			
Average score second set of quizzes (in %)	80.0	66.5	88.3
Average score final exam (in %)	47.8	54.5	62.8
Average score final grade (max is 4.0)	2.45	2.76	2.39

6.2 Discussion

The data presented in Tables 18–20 demonstrate the following:

Improvement by quiz attempt to quiz attempt does not show a consistent pattern. Sometimes averages, and therefore students, show improvements but at other times it does not. In other words, students do not consistently improve on the second attempt and do worse on the third. However, when comparing the first attempt and the last attempt, students who take multiple attempts do improve overall on the quizzes as can be seen from Table 19. Furthermore, the mean first online quiz score for students using only one attempt was quite a bit higher than for students who subsequently repeated a quiz, confirming findings from Kibble (2007).

The performance on online quizzes is correlated with the course grade. This pattern was also found by Metz (2008). However, Metz (2008) found a nearly 1:1 correspondence whereas the data in the current study had a much lower correlation, and it became smaller the less weight the online quizzes had in the course grade.

When comparing the final score achieved on an online quiz, the evidence is inconclusive. In some cases, repeaters score better on online quizzes (see e.g. the summer of 2007 – although the group that repeated some but not all quizzes was worse than the other two groups). In other instances repeaters performed worse, for example, the first half of summer 2006 although the difference is small.

The same type of conflicting evidence was also found when comparing the exam scores for repeaters and non-repeaters of online quizzes.

- Some of the evidence showed that non-repeaters did worse on the exams. For example, during the second half of spring 2006 repeaters of online quizzes scored better on the final exam although those that repeated some but not all quizzes did better than those students who repeated all quizzes. This confirms Kibble (2007) who found for cohorts 3 and 5 that the non-repeaters scored lower on exams than repeaters. This was also found by Smith (2007).
- Some of the evidence showed that there was not much difference between repeaters and non-repeaters. For example the second half of summer 2007. This confirms Kibble (2007) who found for cohorts 2 and 4 that there was no significant difference on exam scores between repeaters and non-repeaters.
- Some of the evidence showed that non-repeaters did better on the exams. For example in the first half of summer 2007. Also, in the first half of summer 2006 the students that repeated all quizzes did worse than the other two groups. This confirms Brothen and Wambach (2001).

When comparing the course grade for repeaters of online quizzes vs. non-repeaters of online quizzes, a similar conclusion can be drawn. The use of online quizzes and multiple attempts on those online quizzes does not lead to better learning as evidenced by higher exam scores or a higher grade for the course.

7 Online quiz impact on the course overall

At the end of summer 2007, it was decided that online quizzes were not significantly contributing to student learning and therefore this element was dropped from the course design. The last research question examines the efficacy of this decision by comparing the summer 2007 quarter (with online quizzes) with the fall 2007 quarter (without online quizzes). This is the third research question. To compare the two quarters, three types of data are compared: the course grade, the mid-term exam grade and the final exam grade. The course grade can be viewed as an indicator for overall learning but a course grade can also include other elements such as attendance/participation. The mid-term exam grades and the final exam grades provide better indicators about what students have learned and provide a more meaningful comparison since these exams are very similar across quarters. Table 21 provides an overview of the results for the two quarters.

Table 21 indicates that the grades received for the two exams as well as for the course overall were slightly higher in summer 2007 compared to the fall 2007 quarter. To determine whether this is a significant difference an ANOVA analysis was carried out, see (Table 22).

The results of the ANOVA analysis indicate that for the mid-term exam as well as for the course grade, the F -value is below the critical value and therefore the hypothesis that the results for summer 2007 and fall 2007 are not different can not be rejected for the mid-term exam and the final grade. In other words, it can not be concluded that the summer 2007 results are statistically different (better in this case) than those in fall 2007. For the final exam, the F -value is higher than the critical value, and hence it can be concluded that the summer results are statistically different, that is, better.

If online quizzes lead to better performance in the course, this should consistently affect both exams, as well as the course grade. Since it can not be concluded that for each of these three items the summer 2007 results are different than those for fall 2007, it is also not possible to conclude that the course was better with online quizzes than without them.

Table 21 Online quiz use and exam and course performance

<i>Quarter</i>	<i>Summer 2007</i>	<i>Fall 2007</i>
Online quizzes used	Yes	No
Number of students included	54	96
Mid-term exam class average	420.0 (75.1)	395.8 (75.2)
Final exam class average	384.8 (87.2)	348.5 (78.2)
Course grade class average	2.6 (0.5)	2.5 (0.5)

Note: The summer quarter followed the same structure as other quarters. That means, the same material is covered, the course is taught in the same format, etc. One difference is that the quarter is shorter than regular quarters, but this is offset by more class hours during a week.

Table 22 ANOVA results for comparing course with online quizzes (summer 2007) with the course without quizzes (fall 2007)

	<i>F-value</i>	<i>P-value</i>	<i>F-critical</i>
Mid-term exam	3.849598	0.051634	3.905058
Final exam	6.838287	0.009846	3.905058
Course grade	2.337324	0.128441	3.905058

8 Conclusion

This paper focused on the use of online quizzes. The purpose of using the online quizzes in the course was to provide a formative assessment instrument that would enable students to learn from quiz feedback and improve their performance. The study took place in a business school setting. This type of setting was not previously covered in the literature on formative online quizzes. Three research questions were posed for this research.

The first research question was: What type of online quiz taking behaviour do students exhibit? It was found that students do not typically follow a prepare-gather feedback-restudy approach. The behavioural patterns found confirmed several observations from previous studies in non-business school settings. Essentially, this pattern is that many students make only one attempt, many students take their first attempt during the last day that quizzes are available, and if a student takes multiple attempts, they are typically taken in a rapid fire fashion. Furthermore, it was found that a grading system that limits the number of attempts and puts time constraints on attempts in addition to subtracting points for simply taking quizzes again rather than studying in between multiple attempts did not change behavioural patterns but instead strengthened previously found patterns.

The second research question was: Do online quizzes help students improve their grade? It was found that students who repeat online quizzes do improve their overall score for the online quiz. It was also found that online quiz scores are related to the grade that students achieve in the course. In general, this confirms the existing literature. However, this is not unexpected since online quiz scores are part of the grade. The less weight online quizzes had, the lower the correlation between online quiz scores and the course grade. Lastly, a distinction was made between repeaters and non-repeaters of online quizzes, and their performance on exams and final grade was compared. This led to inconclusive evidence that is similar to what was noted in the literature. This raises the issue of whether the use of online quizzes is a good mechanism for improving student learning.

The last research question was: Do courses that use online quizzes perform better than courses without online quizzes? The comparison took place by looking at student learning on exams and the final grade earned. The statistical analysis showed that there was no significant difference between the course with online quizzes compared to the course without online quizzes. This type of analysis was not previously conducted in the literature. The conclusion raises questions about the use of online quizzes as an effective course design element to improve student learning.

There are at least two areas worthy of further investigation. The finding that students mostly follow a quiz-to-learn strategy instead of the intended prepare-gather feedback-restudy strategy despite a grading system that was expected to change the behaviour is interesting. It leads to two questions: Can the student approach to studying be changed and if so how? It is possible that cognitive development positions (Perry's stages) have an influence. It is also possible that student engagement in the course or student learning styles in general have an influence. This could not be analysed in this study but might be worthwhile for further investigation.

Secondly, the inconclusive findings related to the comparison of repeaters and non-repeaters of online quizzes and their exam and course performance is intriguing. Brothen and Wambach (2001) noted that the relationship between number of quizzes taken and

quiz performance was not uniform. Although, data has not been presented at the individual student level in this study, a similar remark should be made. What this indicates is that although the evidence is inconclusive when comparing groups of students for example repeaters vs. non-repeaters of online quizzes, it might be possible to find other variables that would lead to better explanations. These other variables are most likely to be found at the individual student characteristics level since results are not uniform.

Acknowledgement

The authors would like to thank Ms. Carlee Marshall for her continued help with the course data analysis.

References

- Anderson, L.W. et al. (Eds.) (2001) *A Taxonomy for Learning, Teaching, and Assessing, A Revision of Bloom's Taxonomy of Educational Objectives* (Abridged ed.). New York: Longman.
- Biggs, J. (1999) 'What the student does: teaching for enhanced learning', *Higher Education Research and Development*, Vol. 18, No. 1, pp.57–75.
- Bloom, B.S. (Ed.) (1987) *Taxonomy of Educational Objectives, The Classification of Educational Goals* (13th Printing). New York: Longman, Inc..
- Brothen, T. and Wambach, C. (2001) 'Effective student use of computerized quizzes', *Teaching of Psychology*, Vol. 28, No. 4, pp.292–294.
- Daniel, D.B. and Broida, J. (2004) 'Using web-based quizzing to improve exam performance: lessons learned', *Teaching of Psychology*, Vol. 31, No. 3, pp.207–208.
- Dopper, S.M. and Sjoer, E. (2004) 'Implementing formative assessment in engineering education: the use of online assessment system Etude', *European Journal of Engineering Education*, Vol. 29, No. 2, pp.259–266.
- Driver, M. (2000) 'Integrating internet-based resources into classroom instruction: an organizational learning approach', *Journal of Business Education*, Vol. 1, pp.14–30.
- Entwistle, N. and Tait, H. (1990) 'Approaches to learning, evaluations of teaching, and preferences for contrasting academic environments', *Higher Education*, Vol. 19, pp.169–194.
- Flamm, P., Hoffman, J.J., Delgadillo, F. and Ewing, B.T. (2008) 'A hybrid approach for teaching introduction to operations management', *Int. J. Information and Operations Management Education*, Vol. 2, No. 3, pp.255–274.
- Francis, V.E. and Schreiber, N. (2008) 'What, no quiz today? An innovative framework for increasing student preparation and participation', *Decision Sciences Journal of Innovative Education*, Vol. 6, No. 1, pp.179–186.
- Gaytan, J. and McEwen, B.C. (2007) 'Effective online instructional and assessment strategies', *The American Journal of Distance Education*, Vol. 21, No. 3, pp.117–132.
- Gose, B. (1997) 'Efforts to curb grade inflation get an F from many critics', *The Chronicle of Higher Education*, Vol. 43, No. 46, pp.A41–A42.
- Greenberger, E., Lessard, J., Chen, C. and Farruggia, S.P. (2008) 'Self-entitled college students: contributions of personality, parenting, and motivational factors', *Journal of Youth Adolescence*, Vol. 37, pp.1193–1204.
- Grimstad, K. and Grabe, M. (2004) 'Are online study questions beneficial?' *Teaching of Psychology*, Vol. 31, No. 2, pp.143–146.

- Gurung, R.A.R. (2003) 'Pedagogical aids and student performance', *Teaching of Psychology*, Vol. 30, No. 2, pp.92–95.
- Guskey, T.R. and Jung, L.A. (2006) 'The challenges of standards-based grading', *Leadership Compass*, Vol. 4, No. 2, pp.6–10.
- Holsendolp, E. (2005) 'Building a 'culture of effort'', *Black Issues in Higher Education*, Vol. 22, No. 5, pp.30–33.
- Kibble, J. (2007) 'Use of unsupervised online quizzes as formative assessment in a medical physiology course: effects of incentives on student participation and performance', *Advances in Physiology Education*, Vol. 31, pp.253–260.
- Lucas, A.F. and Associates (2000) *Leading Academic Change, Essential Roles for Department Chairs*. San Francisco: Jossey-Bass Publishers.
- Maki, W.S. and Maki, R.H. (2001) 'Mastery quizzes on the web: results from a web-based introductory psychology course', *Behavior Research Methods, Instruments, and Computers*, Vol. 33, No. 2, pp.212–216.
- Metz, A.M. (2008) 'The effect of access time on online quiz performance in large biology lecture courses', *Biochemistry and Molecular Biology Education*, Vol. 36, No. 3, pp.196–202.
- Nature (2004) 'Against grade inflation, How to counter declining rigour in US university courses', *Nature*, Vol. 431, No. 7010, p.723.
- Nguyen, D.M., Hsieh, Y.C.J. and Allen, G.D. (2006) 'The impact of web-based assessment and practice on students' mathematics learning attitudes', *Journal of Computers in Mathematics and Science Teaching*, Vol. 25, No. 3, pp.251–279.
- Perry Jr., W.G. (1999) *Forms of Intellectual and Ethical Development in College Years, a Scheme*. San Francisco: Jossey-Bass a Wiley company, Originally published 1968.
- Samuelowicz, K. and Bain, J.D. (2001) 'Revisiting academics beliefs about teaching and learning', *Higher Education*, Vol. 41, pp.299–325.
- Smith, G. (2007) 'How does student performance on formative assessments relate to learning assessed by exams?' *Journal of College Science Teaching*, Vol. 36, No. 7, pp.28–34.
- Taraban, R., Maki, W.S. and Rynearson, K. (1999) 'Measuring study time distributions: implications for designing computer-based courses', *Behavior Research Methods, Instruments and Computers*, Vol. 31, No. 2, pp.263–269.
- Wankat, P.C. and Oreovicz, F.S. (1993) *Teaching Engineering*, Chapter 14. New York: McGraw-Hill.
- Wijekumar, K., Ferguson, L. and Wagoner, D. (2006) 'Problems with assessment validity and reliability in web-based distance learning environments and solutions', *Journal of Educational Multimedia and Hypermedia*, Vol. 15, No. 2, pp.199–215.

Note

¹More information as well as the instrument can be obtained from Dr. William S. Moore, e-mail: wsmoore51@comcast.net.