Issues in Feedback on Student Performance: Timing & Type

Steven A. Yourstone *
S. H. Kraye
Gerald Albaum
University of New Mexico, Albuquerque, New Mexico, U.S.A.

The Operations Management literature has not explored the topic of the timing and type of feedback to students. This important pedagogical topic is explored in the present paper through a literature search covering the years 1926 through 2010, and by an empirical study on the timing of feedback to students in an Operations Management class. The literature search reveals that student learning depends not only on the timing of the feedback, but also the type of feedback, the amount of feedback, the scholastic level of the student, the level of complexity of the learning task, stage of learning, prior knowledge, and whether we are measuring acquisition of knowledge or delayed retention. Using an experimental design, the empirical study looked at comparing student learning (measured by examination scores) using an online grading system with immediate feedback versus delayed paper feedback. Results indicate no significant difference.

*Corresponding Author. E-mail address: yourstone@mgt.unm.edu

I. INTRODUCTION

Researchers have wrestled for many years with the question of feedback type and timing to students. In the present study we examined the literature from the time period 1926 to 2010, and found that the results of numerous studies are mixed. This state of affairs is likely due to the variety of ways in which the studies were conducted and the specific variables that were chosen. The studies differed to some degree on the measurement variable for learning achieved.

There is no simple answer to the question of which type of feedback, and the timing of it, is best. Yet, if there is a best way then it should be identified. After all, the primary concern of educators is learning by students. Any given class and instructor is but one link in the chain of education. Clearly, timing and type of feedback are important matters for each student and instructor. Moreover, each student may react differently to a feedback approach. Like so many other pedagogical techniques, educators must continually search for the technique that is best for the majority of students. Technology also has had a major impact on this issue.

The present study examines feedback in undergraduate Operations Management classes by looking at the effects of immediate versus delayed feedback in quantitative assignments as well as in overall examinations. While the study was conducted in an introductory Operations Management class the students were from all majors and the type of testing is applicable to other classes with quantitative content. A major objective of the study was to determine if quicker online feedback through a course management system was actually better when it comes to a student's learning/understanding of a quantitative skill versus delayed feedback in a more traditional hard copy format. In the online

format feedback can be virtually immediate. However, in the traditional format feedback approach, feedback was received one week following the student submission.

The paper is organized as follows. The next section presents an extensive review of the literature on timing and type of feedback. Also discussed are student forgetting, student's perceptions of feedback, and a discussion of grading quantitative homework assignments. Following the discussion of the literature, the methodology and findings of the empirical study conducted are presented. The last section of the paper concludes the study and suggests areas for future research.

II. LITERATURE REVIEW

2.1. Timing of Feedback

There has been an ongoing debate in the literature about the best timing of feedback to learners. Immediate feedback systems for education have been available since as early as the 1920s. Sidney L. Pressey designed a series of machines that provided automatic testing of intelligence and information (B.F. Skinner, 1958; Pressey, 1926). His machines have been referred to as the "Automatic Teacher." His first machine went into use in 1926. Pressey's automatic teaching machines were simple by today's standards. His teaching machines consisted of a test attached to a rotating drum and keys used to enter the correct multiple choice response. If the learner entered the correct response then the machine advanced to the next question.

The highly cited review by Ammons (1956) on the effects of knowledge of performance and the kinds of information provided in feedback covers important generalizations about knowledge of performance. Ammon's review examined fifty years of the literature regarding knowledge of performance. "In general, research on the problem of knowledge of performance involves determining the effects of giving or withholding various kinds of information about performance during and or varying amounts of time after that performance" (Ammons, 1956, p. 279). The body of research at the time of the Ammon's study had examined several generalizations. One such generalization was that the longer the feedback on performance was delayed the less the effect of the information on the learner. Ammons (1956, p. 289) further argued: it is likely that there is an optimum time of delay for feedback on performance where the effect depends on the task and the stage of learning; if feedback arrived too soon then it might not be used; and delay of knowledge of performance might allow the learner to make a better assessment of his or her performance. Contrary to this, B.F. Skinner (1958, p. 969) stated that delaying the return of examinations after a delay of either hours or days resulted in little measurable modification of student behavior. Skinner (1958) concurs with the work of Pressey (1926) who had emphasized the importance of immediate feedback in the education of students.

The literature identifies additional dimensions along which to analyze the benefits of immediate versus delayed feedback of performance. Bardwell (1981) points out that we must look at learning in terms of acquisition and learning in terms of retention in order to sort out the effects of immediate versus delayed feedback. Acquisition means that the learner made the correct response when the question was first posed. Retention is assessed by the extent to which the learner correctly responds to the same question at a later point in time. Bardwell (1981, p. 4) concludes that delayed feedback facilitates retention. Such a conclusion might seem to run counter to what educators expect.

Kulhavy and Anderson (1972) postulated that during the delay in feedback on performance the learner forgets his or her initial response. And when the delayed feedback is provided to the learner it is postulated that there will be less interference by the initial incorrect response because it has been forgotten. Kulhavy and

Anderson (1972, p. 505) named this the perseveration-interference hypothesis. This hypothesis seeks to explain why there would be better retention with delayed informative feedback. These researchers argue that when learners receive immediate feedback they will experience interference and therefore will be less able to correct their error on tests given at a later point in time. Kulhavy and Anderson (1972, p. 505) concluded that educational theories which assume that immediate feedback is superior have been based on studies where the dependent variable was speed of acquisition rather than the amount learned and the period of time over which the material is retained.

Studies by Sturges (1969, 1972) have concluded that delayed feedback is superior for retention. "The general results of these investigations are that at acquisition or immediate retention there is no significant difference between delayed and immediate feedback but on later retention, delayed feedback is superior" (Sturges, 1972, p. 32). Sturges (1972) postulates that delayed feedback will result in more precise discrimination between the correct and incorrect choices due to the learning of both over the delay period. Based on his and others studies, he reasons that with most forms of feedback learners given immediate feedback do not acquire any significant information at feedback and also immediate application does not aid later retention. At immediate feedback the learner may only be focusing on whether the answer was right or wrong and not on why. Learners need to acquire information about why their answer was right or wrong and not just that it was correct or incorrect. Sturges (1972, p. 43) postulates the importance of the spacing of learning events: "Rather, long-term retention is improved when conditions are such that subjects identify relationships between the to-be-remembered units and other possible alternatives. Perhaps a kind of network is developed in which the correct response is integrated with incorrect alternatives; and long-term retention is better when there is such a network than when subjects have acquired the correct alternative only." The time delay from learner response to the receipt of knowledge of performance seems to create an opportunity for higher order processing of the question, problem, or task. The learner will then perhaps be less interested in whether the answer was correct or incorrect and more attuned to the reasoning behind the correct answer.

In another study Sturges (1978) examined the effect of delay of feedback in computer-assisted testing. He found that retention after a delay of informative feedback was generally superior. Surber and Anderson (1975) found support for the perseveration-interference hypothesis and they found that delayed feedback was significantly better when the initial response was incorrect.

The body of empirical research on feedback timing reveals mixed results. Beck and Lindsey (1979) found that there was no significant difference between feedback given immediately and feedback given one week later. Their study did not distinguish between immediate acquisition and delayed retention. Crooks (1988) found that the precise timing of feedback did not appear to be critical. Kulik and Kulik (1988) found from their meta-analysis of the effects of classroom testing that delayed feedback was superior to immediate feedback only under conditions that were artificial and controlled to some degree. Bangert-Drown et al. (1991) conducted a meta-analysis of 58 effect sizes based on 40 reports. They found a generally weak relationship between feedback and achievement. This result clearly surprised these researchers. Their results indicated that feedback makes a small positive contribution to achievement. In addition they found that control for presearch availability was significantly related to the effect size of feedback on achievement. Presearch is the ability of students to see any feedback prior to formulating their answer (Bangert-Drowns et al., 1991, p. 219). They found that when the students could not look at any feedback about their answers prior to

submitting a response the effect of feedback after the answer had been submitted was significant. When the students could presearch then the effect of feedback after the response was submitted was surprisingly negative. Clearly, finding that feedback could have a negative effect on learning is not what educators would expect nor want. These researchers conclude that presearch availability alters the effect of feedback. "In summary, presearch availability appears to mediate feedback effects, but this apparent influence is confounded with a comparison of feedback effects in programmed instruction to feedback effects in other forms of instruction" (Bangert-Drowns et al. 1991, p. 225). These researchers also found that feedback will have its greatest effect on items that are perceived to be difficult. And feedback was found to have its greatest effect when the student had a high confidence in his or her answer but it turned out to be incorrect. Clariana et al. (2000) reported that the retention of initial test responses was higher for delayed feedback for all levels of item difficulty. They also found this effect to be particularly true with difficult items.

Little (1934) found a significant difference between final exam grades of students who received immediate feedback and those students who did not have this feedback. This researcher discovered that students given immediate feedback and the opportunity to correct deficiencies by makeup tests showed higher final exam scores than those students who did not have the same advantage. The students using Pressey's automated teaching machine answered each question until they achieved the correct response. More interestingly, in Little's study it was found that drill and immediate feedback was most effective for students in the lower half of the scholastic distribution. The entire group of students provided immediate feedback moved up in performance overall. Little (1934, p. 49) concluded that mechanical testing devices were valuable in the classroom due to their immediate feedback and speed.

Shute (2008) concludes that the effect of the timing of feedback is unclear in the body of research she examined, and reasons that immediate feedback and delayed feedback have both positive and negative learning effects. Immediate feedback can provide motivation to practice and an immediate association between input and response. The negative side of immediate feedback might be less careful processing of the question or task. Shute argues that delayed feedback can have a positive effect of encouraging learners to engage in active cognitive processing. And on the negative side delayed feedback can frustrate struggling and less motivated learners (Shute 2008, p. 166).

2.2. Type of Feedback

The effects of feedback also depend on the type of feedback provided. Mason and Bruning (2001, p.1) examined feedback in computer-based instruction and identified eight levels of feedback: (1) no-feedback, (2) knowledge-of-response, (3) knowledge-of-correct-response, (4) answer-until-correct, (5) topic-contingent, (6) response-contingent, (7) bug-related, and (8) attribute-isolation . In addition, they identified five variables which may influence the effectiveness of feedback: (1) student achievement, (2) task complexity, (3) timing of feedback, (4) prior knowledge, and (5) learner control. These researchers found that the results based on research on the effects of feedback elaboration are contradictory. They examined several studies on levels of feedback to identify differences that might have affected the study results, and found methodological concerns in some of the studies. They reported that the nature of the topic and the skill being tested also limited the generalizability of the results.

Whyte et al. (1995) studied the effect of feedback as the level of feedback varied. They found that the largest gains in learning with computer based instruction resulted from the most elaborate levels of feedback. The highest

level of feedback as well as gains in learning resulted from the knowledge-of-response plus response-contingent feedback. The knowledge-of-response feedback is the simplest form of feedback, in which the learner is provided a score with no details on individual test items. In contrast, the response-contingent feedback provides the learner with knowledge of the correct response and specific feedback about why the incorrect answer is not correct and why the correct answer is correct. The finding that the most elaborate feedback results in the largest gains in learning seems to be an obvious result. It is something that many of the online course management systems are beginning to offer to some degree. They do offer knowledge-of-response and answer-until-correct. But they do not offer response-contingent feedback. This has always been an advantage of work that is graded by hand with elaborate feedback provided to the student. Persky and Pollack (2008) found that pharmacy students given examinations with immediate feedback and answer-until-correct grading option did not outperform those given traditional delayed feedback with single attempt grading. Mason and Bruning (2001, p. 6) conclude that the body of research does not isolate a single type or level of feedback. But they do identify a trend towards increased learning in response to more elaborate feedback.

Research has revealed that the type of feedback is moderated by the ability level of the student. Clariana (1990) found that low ability students had higher learning outcomes as a result of receiving knowledge-of-response feedback compared with answer-until-correct feedback.

Epstein et al. (2001) studied two classes of undergraduate introductory psychology. They found that students receiving immediate feedback on unit exams during the term then correctly answered more questions on the final exam where the questions were repeated from the unit exams. The form of the unit exams in the immediate feedback section was answer-until-correct. And the students in that same section answered each question correctly before moving on to the next question. The other section of the class used delayed feedback unit exams. Epstein et al. (2001) mixed delayed versus immediate feedback with the answer-until-correct testing format. This makes it difficult to isolate the effect of immediate versus delayed feedback. And they conclude that immediate feedback aided delayed retention. The delay point in time was at the end of the academic term.

In a study on the role of feedback in academic testing Dihoff et al. (2003) found that immediate feedback combined with answer-until-correct responses leads to enhanced retention over the course of an academic semester. These researchers note that there is agreement in the field that learning is enhanced by feedback. But they find that there is little agreement in terms of what type of feedback is more effective. Dihoff et al. (2003, p.13) conclude that the greatest retention, greatest student confidence, and the greatest accuracy came from the provision of immediate feedback coupled with answer-until-correct. These researchers postulate that the corrective information provided by immediate feedback might prevent the learner from committing to a response. They argue that the learner will then be in a state of disequilibrium which will be resolved by acquiring the correct response. This conclusion appears to be in conflict with those that support the perseveration-interference hypothesis (Sturges, 1978; Surber and Anderson, 1975). The study by Peeck and Tillman (1979) also does not support the perseveration-interference hypothesis. Not only did the studies surveyed here show mixed results they are also contradictory in their explanations of the outcomes.

Butler et al. (2007) designed a study which included two experiments to isolate the effects of answer-until-correct and immediate versus delayed feedback on long-term retention. Interestingly these authors point out that the benefit in answer-until-correct versus standard feedback is that the student must discover the

answer rather than to just be given the correct answer in the standard feedback format. The difference between the two approaches to feedback is essentially active versus passive processing. In answer-until-correct feedback the student actively seeks the correct answer. In the standard feedback situation the student passively receives the answer. And in the most limited type of feedback the student only receives the information that the answer is correct or incorrect. An interesting result of the Butler et al. (2007) study was that in the first of two experiments the researchers found that the answer-until-correct alternative did not produce a significant benefit over standard feedback of correct/incorrect. This is contrary to the findings of previous studies. This result is likely due to the design of the study which was intended to untangle the effects of answer-until-correct and immediate versus delayed feedback. Butler et al. (2007) also performed a second experiment that lengthened the time delay of the feedback from one day used in the first experiment to one week. In the second experiment they found that the magnitude of the advantage of delayed feedback was even greater. This can be explained by the spacing of learning events to one week as opposed to one day. Spacing of learning events as opposed to massing of learning events at the same point in time provides the learner the opportunity for learning at several points in time. In the second experiment, as in their first, they found no significant difference between answer-until-correct and standard feedback. Butler et al. (2007) found in both experiments that delayed feedback produced better long-term retention than immediate feedback. These researchers effectively isolated the effects of type and timing of feedback. Most prior studies had not done so.

Shute (2008) performed an extensive literature review on formative feedback. Shute (2008,p 154) identified dozens of feedback types including accuracy of solution, topic contingent, response contingent, attribute-isolation worked examples, hints, and partial solutions and concludes that different studies reported opposing conclusions about the same feedback variable. Due to the makeup of the studies examined by Shute it is possible that interactions among the subsets of variables chosen masked individual variable effects and led to disparate conclusions regarding the same variables. Shute observed that the specific linkages between feedback and learning are still unclear even with the large body of research on the subject and suggests that the most effective feedback is a combination of correct/incorrect and elaboration to guide the learner towards the correct answer. Shute (2008, p. 16) provides a comprehensive listing of feedback types. Listed by complexity, these are: (1) No feedback, (2) Verification, (3) Correct response, (4) Try again, (5) Error flagging, (6) Elaborated, (7) Attribute isolation, (8) Topic contingent, (9) Response contingent, (10) Hints/cues/prompts, (11) Bugs/ misconceptions, and (12) Informative feedback. Clearly when researchers discuss immediate versus delayed feedback it is necessary to identify these other dimensions of feedback. Shute (2008, p. 182) concludes "As evidenced throughout, there is no "best" type of formative feedback for all learners and learning outcomes."

2.3. Forgetting What They Had Learned

All of the literature discussed above dealt with measuring learning but not forgetting over a period of time following a class. Moreover, these prior studies measured retention of learning during and at the end of an academic term. The ability to apply topics learned in a class at points in the future requires that a student remember what was learned into the future. Bacon and Stewart (2006) examined how fast students forget what they learned in a consumer behavior class. Bacon and Stewart examined retention over the very long term ranging from 8 to 101 weeks. They found that most knowledge was lost within two years. They focused on the steepness of the retention curve for student knowledge of consumer behavior. And they discuss seven recommendations for

reducing the amount of forgetting that takes place: "(1) develop a pedagogy that requires deep learning, (2) sacrifice breadth for depth, (3) require that students take a course's prerequisites immediately before a course, (4) focus course content on concepts and tools that students will encounter in their first jobs, (5) use cumulative exams, (6) assume a broader approach to teaching evaluation, and (7) implement learning assessments across a wider time frame" (Bacon & Stewart, 2006, p. 189).

When we consider the work of Bacon and Stewart (2006, p. 182) together with the work of Kulhavy and Anderson (1972, p. 505) a linkage emerges. Kulhavy and Anderson (1972) postulated the perseveration-interference hypothesis. This hypothesis is concerned with the effect of the first response to a question and how it interferes with the correct response if the feedback is given soon afterwards. Over time it is postulated the first response will cease to interfere with giving the correct response to the same question at a later point in time. Focusing on Marketing classes, Bacon and Stewart (2006, p. 183) note two interference effects being created by two sets of memories - proactive and retroactive. Proactive interference may occur when the two sets of memories are created by two different consumer behavior frameworks being taught, for example, in the basic marketing principles class versus the consumer behavior class. Retroactive interference may occur when two different approaches to the same consumer behavior model are taught within the same class. The mechanisms for proactive interference, retroactive interference, and perseveration-interference might all be similar in their causes and in their effects on delayed retention and eventual forgetting. Although this research was done within the consumer behavior area, the findings might very well apply to other areas such as Operations Management, Finance, Accounting, etc.

2.4. Perceptions of Feedback by Students

Ackerman and Gross (2010) highlight the importance of feedback in the education process for marketing students yet point out that the marketing education literature has given it little attention. The findings of their study suggest that when an instructor provides a lot of feedback on an assignment students tend to receive it negatively. These researchers discuss the emotional component of student feedback. Some students respond better to criticism than others. When students receive high levels of feedback they tend to view the professor as being more critical of them. The academic maturity of the student clearly plays an important role in the effectiveness of the type and timing of feedback. A great deal of effort is invested by instructors in providing feedback. But that feedback may not be received effectively if the given emotional state of the student is not one that is receptive. There may be a negativity effect (Ackerman and Gross, 2010, p.174.) They postulate that students might view constructive feedback as negative if it is directed towards that particular student, which may lead to a negative impression and resulting distancing from the instructor. This should not be interpreted to mean that instructors should only give feedback that is interpreted as praise. Prior research has indicated that feedback given in the form of praise by itself will not lead to higher learning outcomes (Hattie & Timperly, 2007; Wilkinson, 1981; Kluger & DeNisi, 1998, Brophy, 1981; Bond at al.1981). Praise given to students does little to guide the student towards an understanding of the material. Praise by itself contains little information that the students may utilize to improve their understanding and can even distract from that objective. Students might not accept feedback that is at odds with their self-perception of their academic abilities. Ackerman and Gross (2010) suggest that if an instructor wished to provide a lot of feedback then the student should be allowed to revise and resubmit. This suggestion is consistent with the findings in the literature that answer-until-correct

can be effective in enhancing student retention of learning based on examinations.

2.5. Grading Quantitative Homework Assignments

Only one paper was found that was relevant to the issue of grading quantitative Operations Management homework assignments. There may have been other studies on this topic, but they were not prominent in the search. Peters et al. (2002) studied the effect of required quantitative assignments on exam performance. Their study included 330 students, involving 13 classes of introductory Operations Management, over three semesters. Their study examined the effect of grading assigned quantitative homework versus just assigning the homework. Their study led to the conclusion that students who are required to turn in quantitative homework for grading actually underperformed those students who were not required to turn in assigned quantitative homework for grading. Interestingly they found a negative association between graded required quantitative homework and exam performance. In all classes the students were informed that the types of quantitative homework problems that they were working would also appear on their exams. Their findings might indicate that students who are graded on assigned quantitative homework, and know that the same types of problems will appear on exams, are not any more or less motivated to learn due to the incentive of grading the homework. These researchers further postulated that it is possible that the negative association of graded quantitative homework and exam performance may be due to the students focusing relatively too heavily on the quantitative topics to the detriment of time spent on the non quantitative topics. They also found that there was a positive association between a student's GPA and exam performance, and that this relationship held regardless of whether or not the assigned quantitative homework was graded. The better students, as measured by GPA, tend to be more motivated to learn the material without the extrinsic motivation of a grade being given to their quantitative assignment.

III. EMPIRICAL STUDY

The focus of the present study was to measure the quantitative differences in learning outcomes between classes that were taught using online grading systems with immediate feedback and multiple numbers of attempts on quantitative assignments versus classes that were provided hardcopy assignments giving students only a single attempt to solve the problem and were hand graded by a professor and returned to the students a week later. This study also examined the effect of concepts quizzes offered online with immediate feedback versus those same types of quizzes offered in hardcopy with delayed grading feedback. Using an online grading system makes it possible to give students additional attempts on the same computational questions but with different data on each attempt. As mentioned earlier, the following research question is the basis of the current study:

Research Question: Do students learn more, as measured by examination scores, using online grading systems with immediate feedback than those students that receive delayed paper feedback?

3.1 Methodology

Research Design

The study was conducted by testing two upper division classes of Operations Management involving a total sample of 90 students (45 students in each class), using an experimental design. Students in each class section in this study sample attended 75 minute classes two days a week during the sixteen-week semester. The same instructor taught both

classes. Both classes used the same textbook, Heizer & Render (2005). One section was taught using the textbook and *PH Grade Assist* for assignments and quizzes. The other class used the same textbook, but assignments and quizzes were completed on paper and the assignments were hand graded. The choice of which class was to receive which approach was done randomly, resulting in a simple completely randomized experimental design.

Each of the two classes was given the same two-part entrance exam over the course material that would be presented during the semester. One part of the entrance examination focused on quantitative knowledge that required some mathematical computations (19 questions). The other part of the entrance exam focused on concepts and definitions (68 questions). The entrance exam was given to all students in the study as a traditional paper exam.

The examinations were given over two class periods and students had one hour and fifteen minutes to complete each exam. Students were not given the results of their entrance exam for two weeks and were only given their raw scores on both examinations. The examinations were not returned to the students. All questions on both of the entrance examinations were multiple-choice with four possible answers.

In addition to the entrance exam and quizzes discussed above, a midterm and final exam were administered to both classes. The midterm exam was 75 minutes long and consisted of quantitative and concept questions and was given as a paper exam to both classes. The midterm exam was a multiple choice format. Students were given copies of the midterm exam to use for study purposes. The midterm exam contained similar concept and quantitative questions to the entrance exam, but with different values and answer choices.

The final exam consisted of 51 concept and 16 quantitative questions and was also a two part paper exam for both classes. The final exam was a two hour in class exam and was composed of the same questions that were given on the entrance examinations, but arranged in a different order. The purpose for giving the same final and entrance exam was to have a benchmark to measure quantitatively student learning in and across both of the classes. Students were not given a copy of the entrance exam or told that the final exam would consist of the same questions as were on the entrance exam.

In addition to the entrance, midterm, and final exams each of the two classes were given 11 problem assignments and 13 multiple-choice quizzes averaging 10 questions per quiz. PH Grade Assist was used in one class section for the 11 problem assignments and the 13 multiple choice quizzes but not on the entrance, midterm or final examinations.

Online Grading with *PH Grade Assist*

As mentioned above, one of the classes used *PH Grade Assist* for assignments and quizzes during the semester, and the other class was given assignments and quizzes in a hard copy format. Each class period consisted of an interactive lecture on an Operations Management subject. During the last 15 to 20 minutes of 13 of the 32 classes, students using their laptop computer were directed to go on line and take a designated quiz that was posted on a server. After all the students had completed the quiz, the Instructor would tell the students to click the "submit" button and they could see their score on the quiz and which questions they missed with the correct answer.

Problem assignments were posted online for students to take during non-class time. *PH Grade Assist* allowed a student to work the problem and submit an answer. Once the student's answer was posted, *PH Grade Assist* would show the student the correct answer. If the student did not receive full credit on all questions, and additional attempts were still possible, then *PH Grade Assist* allowed the student to attempt the problem again but with different data. *PH Grade Assist* showed the student the number of points earned out of the

total possible on each assignment. The student could see which problems on the assignment received full credit or partial credit assigned automatically by the system. The student could submit another attempt on an assignment to *PH Grade Assist* for three to five times (depending on what the Instructor set up, i.e., for more difficult problems students were given 5 attempts and for simpler ones students were given 3 attempts). Scores were recorded on the *PH Grade Assist* website and made available to the Instructor for review and download as a spreadsheet.

Students were encouraged to ask questions about the online problem assignments they missed and similarly for the in class hardcopy assignments in the traditional class section.

How Traditional Classes Operated

Students in the traditional class section were given the same quizzes and assignments as the online classes, but in a traditional paper format during the last 15 to 20 minutes of the class period. The traditional section was always given the same total time to finish a quiz as the online classes had. The traditional quizzes were machine graded and returned the following week. If a student taking the paper quiz had a question about an answer to a question on the quiz, it was answered at the time the quizzes were returned. The quantitative assignments in the traditional section were submitted in hardcopy, graded by the instructor, and returned within one week. However the traditional class section did not have multiple attempts at the same assignment problems with new data each time.

3.2. Findings

As previously mentioned, the research question being investigated is: Do students learn more, as measured by examination scores, using online grading systems with immediate feedback than those students that receive delayed paper feedback?

Entrance and final examination scores were the basic measurements used in the analysis. As previously mentioned, the examinations contained identical questions, but the order was randomized between the classes. An analysis was first made of the difference in final and entrance exam scores (See Table 1). There was a significant difference (*p*=.01) in both the immediate feedback and delayed feedback experimental groups. Thus, overall the students did improve upon their knowledge of introductory Operations Management.

TABLE 1. LEARNING OUTCOME ANALYSIS

		Examination Score		
Experimental Group	N	Entrance	Final	*p*
Immediate Feedback	45	31.7	42.3	0.01
Delayed Feedback	45	31.4	40.6	0.01

Table 2 shows that there was no significant difference found between the class section with immediate feedback and the section with delayed feedback on the entrance total exam scores, as well as quantitative and qualitative exam scores. Table 2 also shows that there was no significant difference on the final examination scores on the same three measures. We conclude that the answer to our research question is "No." Students did not learn more, as measured by examination scores, using online grading systems with immediate feedback than those students that receive delayed paper feedback. This result is even more interesting given that the students who were provided

immediate feedback were also allowed to have multiple attempts on each quantitative problem.

Multiple regression analysis was used to assess the relationship between the individual level differences in entrance and final examination scores (dependent variable) and the independent variables (1) total correct on entrance exam, (2) total correct on assignments, and (3) total correct on quizzes. Since the dependent variable is a measure of individual outcome or learning, the independent variables are related. The resulting regression statistics for delayed (online) and immediate (traditional) feedback are shown in Table 3. Significant predictors for both class sections are entrance examination score and performance on quizzes. The signs of the beta coefficients are as one would expect. That is, the higher the entrance exam score the lower the individual difference between entrance and final exam (Traditional β=-0.828/Online β=-0.616). Interestingly, performance on assignments was not significant (Traditional p=0.278/Online p=0.724). And the higher the quiz score the higher the individual difference between entrance and final exam (Traditional β=0.199/Online β=0.154). The adjusted R square values for both groups were similar. The adjusted R square value for the traditional section was 0.4607 and for the online section it was 0.5021. Both explain about 50% of the variation in the dependent variable. This suggests that there may be other explanatory variables (e.g., GPA) to include in future models to increase the explanatory power of the model.

TABLE 2. EXAMINATION SCORES BY EXPERIMENTAL GROUP

Examinations	Immediate (N=45)	Delayed (N=45)	p
Entrance Examination			
Total Score	31.7	31.40	0.84
Quantitative Score	7.8	7.70	0.91
Qualitative Score	23.8	23.70	0.54
Final Examination			
Total Score	42.3	40.6	0.21
Quantitative Score	11.5	10.6	0.27
Qualitative Score	30.9	29.6	0.20

TABLE 3. REGRESSION MODEL RESULTS

	Delayed[a b]		Immediate[cd]	
	β	p	β	p
Constant	25.32	0.001	19.07	0.001
Total Correct on Entrance Exam	-0.828	0.001	-0.616	0.001
Total Correct on Assignments	-1.098	0.278	-0.009	0.724
Total Correct on Quizzes	0.199	0.045	0.154	0.004

[a] F=13.53, p=0.001
[b] *Adjusted R Square= .4607*
[c] F=15.79, p=0.001
[d] *Adjusted R Square=.5021*

The analysis shows that there was no significant difference in the total exam scores or the quantitative and qualitative section scores on the entrance exam. There was also no significant difference in the total final exam or the quantitative or qualitative section scores between immediate feedback and delayed feedback. There was no significant difference in the change in scores from entrance to final exam between the traditional hardcopy section and the online grading section.

IV. CONCLUSION

The results of our empirical study revealed that there was no significant difference in learning outcomes due to the timing of the feedback. This result may be due to the influence of other variables that may mask a true difference. Or it may actually be the case that the timing of the feedback on quantitative assignments does not have any significant effect on student learning. The time delay period was not specifically studied in levels. The time delay can be set at immediate, one hour, several days, or any delay that one wishes. So it is possible that the time delay variable is more complex than immediate versus some fixed delay. This clearly represents an opportunity for future research. The feedback in online course management systems can also be modified such that the student can check for correct/incorrect answers prior to submitting the assignment for grading. That feature was not enabled in the current study.

Much has been written on the timing and type of feedback. The literature discussion that we have provided leads us to the conclusion that the effects of different types of feedback are not clearly identified due to the mixed results of prior studies when it comes to the timing of feedback.

Numerous research studies were noted in the literature survey, which reveals that the results of studies on the timing of feedback are mixed. And these studies include a wide variety of fields such as operations management, pharmacy, psychology, military training, statistics, and marketing. It is difficult to say with any statistical level of confidence that immediate feedback differs significantly from delayed feedback when measuring the same learning outcomes for students as measured by examination scores. While many studies in the literature conclude that quicker feedback is better, there are an equal number of other studies that conclude that delayed feedback is better. The literature identifies other variables that when interacting with the timing of feedback may alter the level of learning outcomes. While the timing of feedback is important the type of feedback is also a key. Student scholastic level (overall GPA), complexity of the task or problem, the level of feedback, stage of learning, gender, spacing of learning events, student self-perception, student emotional state, whether the first response was correct or incorrect, and other variables can interact with and affect the impact of the timing of feedback. One obvious additional variable in the present study was the additional attempts provided to the immediate feedback section for the quantitative assignments. That said, one would have expected that the immediate feedback section would have outperformed the section with delayed feedback. That did not occur. In fact, the present study did not find a significant difference in learning outcomes in the Operations Management classes studied.

There may also be an effect due to the experience that students may have with online grading and course management systems. And some students may not learn best in such an environment. Individually, instructors also vary in their desire, experience, and ability in the use of online course management systems. Furthermore the online course management systems also vary across vendors in terms of their functionality, ease of use, and sophistication.

We can glean some useful general guidelines for operations management

instructors from the literature survey and the empirical study. First, quicker feedback does not necessarily benefit instructors or their students. In the present empirical study it was concluded that immediate feedback does not differ from delayed feedback. As one can clearly see from the literature survey student learning is more complex than that. For higher level learners the evidence gleaned from the literature seems to suggest that delayed feedback allows for additional high level processing. The nature of the feedback in combination with the timing of the feedback should also be considered. Context specific feedback is more tailored to the individual student and can provide better individualized guidance as opposed to correct/incorrect type of feedback. Written comments can also negatively affect the motivation of the student to continue to learn as shown by Ackerman and Gross (2010). Perhaps a less is more strategy in terms of written feedback will run less of a risk of shutting down some students. Online course management systems allow for multiple attempts to correctly answer the same quantitative questions. The literature seems to support the conclusion that answer-until-correct can be beneficial to the student.

While instructors recognize the importance of feedback to their students it is not entirely clear based on the body of research what is the best type of feedback and what is the best timing of feedback. So clearly there exists a rich opportunity for future research on these topics. And the contribution will be to inform instructors of the value of different types and timing of feedback and the best uses of their time in providing feedback to students. The same entrance and exit exams were used to measure learning. This allowed for use of a matched pair analysis. It would be interesting in future research to see if using different but similar content exams for entrance and exit would reveal the same or different results.

Other areas for future research are an empirical analysis of the different feedback types. Moreover, since the present study deals with only one particular aspect of timing, (the student performance results based on immediate feedback vs. delayed feedback) there is a need to analyze other aspects of feedback timing, for example, an optimal duration of the delayed feedback, e.g., should it be one day, one week, two weeks, etc. The Operations Management literature has not carefully examined the important issues on student performance as affected by the timing and type of feedback.

V. REFERENCES

Ackerman, D.S., & B.L. Gross, "Instructor Feedback: How Much Do Students Really Want?", *Journal of Marketing Education*, 32(2), 2010, 172-181.

Ammons, R.B., "Effects of Knowledge of Performance: A Survey and Tentative Theoretical Formulation", *The Journal of General Psychology*, 54, 1956, 279-299.

Bacon, R.D., & K.A. Stewart, "How Fast Do Students Forget What They Learn in Consumer Behavior? A Longitudinal Study", *Journal of Marketing Education*, 28 (3), 2006, 181-192.

Bardwell. R., "Feedback, How Does it Function?", *The Journal of Experimental Education*, 50(1), 1981, 4-9.

Bangert- Drowns, R.L., Kulik, L,C., Kulik, J.A., & Morgan, M.T., "The Instructional Effect of Feedback in Test-Like Events", *Review of Educational Research*, 61(2), 1991, 213-238.

Beck, F.W., & Lindsey, J.D., "Effects of Immediate Information Feedback and Delayed Information Feedback on Delayed Retention", *The Journal of Educational Research*, 72(5), 1979, 283-284.

Bond, L., Smith, R., Baker,W.K., & Hattie, J.A. "Certification System of the National Board for Professional Teaching Standards: A Construct and Consequential Validity Study", 2000, Washington, DC: National Board for Professional Teaching Standards.

Brophy, J., "Teacher Praise: A Functional Analysis", *Review of Educational Research*, 51, 1981, 5-32.

Butler, A.C., Karpicke, J.D., & Roediger III, H.L. "The Effect of Type and Timing of Feedback on Learning From Multiple-Choice Tests", *Journal of Experimental Psychology, p. Applied*, 13(4), 2007, 273-281.

Clariana, R.B. "A Comparison of Answer-Until Correct Feedback and Knowledge-of-Correct Response Feedback Under Two Conditions of Contextualization", *Journal of Computer-Based Instruction* 17(4), 1990, 125-129.

Clariana, R.B., Wagner. D., & Roher Murphy, L.C., "Applying a Connectionist Description of Feedback Timing", *Educational Technology Research and Development*, 48(3), 2000, 5-21.

Crooks, T.J, "The Impact of Classroom Evaluation Practices on Students", *Review of Educational Research*, 58(4), 1988, 438-481.

Dihoff, R.E., Brosvic, G.M., & Epstein, M.L., "The Role of Feedback During Academic Testing: The Delay Retention Effect Revisited", *The Psychological Record*, 53(4), 2003, 533-548.

Epstein, M.L, Epstein, B.B., and Brosvic, G.M., "Immediate Feedback During Academic Testing", *Psychological Reports*. 88, 2001, 889-894.

Hattie, J, & Timperley, H., "The Power of Feedback", *Review of Educational Research*, 77(1), 2007, 81-112.

Heizer. J & Render. B., *Operations Management*, Flexible, 8th Ed. Prentice Hall, Upper Saddle River, NJ, 2005

Kluger, A. N., & DeNisi, A., "The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory", *Psychological Bulletin*, 119(2), 1996, 254-284.

Kulhavy, R.W., & Anderson, R.C., "Delay-Retention Effect with Multiple Choice Tests, *Journal of Education Psychology*, 63, 1972, 505-512.

Kulik, J.A., & Kulik, C-L. C., "Timing of Feedback and Verbal Learning", *Review of Educational Research*, 58, 1988, 79-97.

Little, J.K., "Results of Use of Machines for Testing and for Drill Upon Learning in Educational Psychology", *Journal of Experimental Education*. 3, 1934, 45-49.

Mason, B.J., & Bruning, R., "Providing feedback in computer-based Instruction: What the research tells us," Center for Instructional Innovation. University of Nebraska-Lincoln; 14. http//dwb.unl.edu/Edit/MB/MasonBruning.html (accessed September 21, 2010).

PH Grade Assist. Pearson Education, Inc. Upper Saddle River, New Jersey.

Peeck, J., & Tillman, H.H., "Learning from Feedback: Comparison of Two Feedback Procedures in a Classroom Setting", *Perceptual and Motor Skills*. 48, 1979, 351-354.

Persky A.M., and G. M. Pollack. "Using Answer-Until-Correct Examinations to Provide Immediate Feedback to Students in a Pharmacokinetics Course", *American Journal of Pharmaceutical Education*. (72)4, 2008, 83.

Peters, M., Kethley, B., & Bullington, K., "The Relationship Between Homework and Performance in an Introductory Operations Management Course", *Journal of Education for Business*, 2002, 340-344.

Pressey, S.L., *School and Society*, 23, 1926, 586.

Shute, V., "Focus on Formative Feedback", *Review of Educational Research*, 78, 2008, 153-189.

Skinner, B.F., "Teaching Machines", *Science*, 128, 24 October 1958. Number 3330, 969-977.

Sturges, P.T., "Verbal Retention as a Function of the Informativeness and Delay of Informative Feedback", *Journal of Educational Psychology*, 60, 1969, 11-14.

Sturges, P.T., "Information Delay and Retention: Effect of Information in Feedback and Tests", *Journal of Educational Psychology*, 63, 1972, 32-43.

Sturges, P.T., "Delay of Informative Feedback in Computer-Assisted Testing", *Journal of Educational Psychology*, 70(3), 1978, 378-387.

Surber, J.R., & Anderson, R.C., "Delay-Retention Effect in Natural Classroom Settings", *Journal of Educational Psychology*, 67(2), 1975, 170-173.

Whyte, M.M., Karolick, D.M., Neilsen, M.C., Elder, G.D., & Hawley, W.T., "Cognitive Styles and Feedback in Computer-Assisted Instruction", *Journal of Educational Computing Research*, 12(2), 1995, 195-203.

Wilkinson, S.S., "The Relationships of Teacher Praise and Student Achievement: A Meta-Analysis of Selected Research", *Dissertation Abstracts International*, 41 (9-A), 1981, 3998.