

**THE EFFECTIVENESS OF ELABORATIVE FEEDBACK ON ACADEMIC
PERFORMANCE IN A WEB-BASED INTRODUCTORY PHARMACOLOGY
ELECTIVE**

Yenna Salamonson RN, BSc, Grad Dip Nsg Stds (Ed), MA, PhD
Senior Lecturer
School of Nursing
University of Western Sydney
Sydney, Australia

Bronwyn Everett RN, BAppSc(Nurs), MSc(Hons)
PhD Candidate
School of Nursing
University of Western Sydney
Sydney, Australia

Jane Koch RN, RNT, MA
Lecturer
School of Nursing
University of Western Sydney
Sydney, Australia

Steven A. Frost RN, BN, MPH
Lecturer
School of Nursing
University of Western Sydney
Sydney, Australia

Corresponding Author:

Yenna Salamonson RN, BSc, Grad Dip Nsg Stds (Ed), MA, PhD

Senior Lecturer

School of Nursing, College of Health & Science

University of Western Sydney

Campbelltown Campus, Building 7

Locked Bag 1797

Penrith South. DC 1797

New South Wales, Australia

Phone: 02 4620 3322

Fax: 02 4625 3161

International: Replace 02 with 612 after your international access code

Email: y.salamonson@uws.edu.au

ABSTRACT

Response feedback in web-based multiple-choice questions can assist students to gauge their understanding of learning material. This study compared the relative effectiveness of ‘verification only’ and ‘elaborative’ response feedback, in a web-based introductory pharmacology course. Data collected on 678 undergraduate students showed elaborative feedback was more effective than verification only feedback, reflected by the higher scores on the summative assessment. While this difference was significant in the two lower GPA groups ($P=0.002$ and $P=0.001$ respectively), the gap was not significant ($P=0.186$) among students with high GPAs, suggesting a greater need for providing elaborative feedback, particularly for students with lower GPAs.

INTRODUCTION

Web-based instruction or online teaching and learning, whereby courses are delivered fully online, is now an integral part of teaching and learning in higher education (Tallent-Runnels et al., 2006). This has been made possible with the Internet, which has the potential to deliver *en masse* highly engaging and meaningful learning materials to the student’s desktop (or laptop), at any time of the day or night. In line with the expeditious growth of web-based courses, the body of literature related to online instruction has expanded rapidly; however, the focus has been on the development and skills required for successful web-based teaching and learning, with lesser emphasis on the effectiveness and the learning outcomes of online courses (Cuellar, 2002).

The use of feedback to promote online learning and improve learning outcomes is often a neglected attribute of web-based instruction (Clariana, Ross, & Morrison, 1991). Feedback can be computer-based and automatically generated, without having to rely on an online instructor. This feedback can be incorporated into multiple-choice questions in the form of self-test activity, to review and revise learning material. Feedback can also be provided through online assessment programs such as online quizzes. These activities provide opportunities for students to actively interact with the learning materials; they also enable the students to gauge the extent to which they have understood the learning material (Chang, 2007).

Although self-test activities that consist of multiple-choice questions have been shown to improve students' retention of materials tested (Roediger & Karpicke, 2006), multiple choice questions expose students to erroneous information, also known as multiple-choice lures, resulting in the learning of false facts as a consequence of faulty reasoning (Marsh, Roediger III, Bjork, & Bjork, 2007). This potential for acquiring misinformation can be circumvented by providing feedback, which highlights erroneous thinking and supplies the correct response (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991).

Two commonly used feedback formats in web-based instruction are verification only feedback and elaborative feedback (Kulhavy & Stock, 1989). The 'answer-until-correct' (AUC) feedback, whereby the student is required to click until the correct answer is selected, is a form of verification only feedback, and does not provide further information as to why a particular answer is correct or incorrect (Mandernach, 2005). In contrast, elaborative or explanatory feedback identifies the correct response and provides relevant cues or explanation for the correct or incorrect answer (Morrison, Ross, Gopalakrishnan, & Casey, 1995).

According to Mason & Bruning (2001), the view shared by most researchers is that elaborative feedback is more effective than verification only because it strengthens correct responses and promotes the retention of new information. However, the writing of quality multiple choice questions with elaborative feedback requires some expertise and, more importantly, considerable additional staff time (Clariana et al., 1991). Although several studies have shown that elaborative feedback is more effective in enhancing learning than simple verification only feedback (Bangert-Drowns et al., 1991; Pridemore & Klein, 1995; Roper, 1977; Whyte, Karolick, Nielsen, & Elder, 1995), some studies failed to demonstrate the added benefits of computer-based elaborative feedback on student learning (Mandernach, 2005; Mory, 1994; Park & Gittelman, 1992). One explanation for these contradictory findings could be the differences in study designs, as there are few well-controlled empirical studies that have manipulated the elaborative feedback component in a systematic manner (Kulhavy & Stock, 1989).

Taking into consideration that more evidence is still required to quell the debate regarding the added benefits of elaborative feedback in enhancing learning, this current study aimed to compare students' performance in their summative assessment and the types of feedback provided in the web-based instruction. The study also sought to examine the relationship between students' characteristics and the type of feedback (verification only or elaboration) that was more beneficial for them.

METHODS

Study Subjects

The study subjects were undergraduate students enrolled in an introductory pharmacology elective offered by the College of Health and Science at the University of Western Sydney, Australia. Most who enrolled in this elective were second year students, midway into their study program.

This study was reviewed and approved by the university ethics committee for human subjects. Students' data were linked to their academic performance and grade point average (GPA) at the end of the semester; however students' names and identity numbers were deleted to maintain their anonymity and confidentiality.

Study design

The study was conducted over two academic years. To maintain consistency in online quizzes, elaborative feedback was provided for all questions while verification only feedback (i.e. Answer-until-correct [AUC]) was used for all self-test questions. The same sighted question set used as online quizzes in Year 1 was converted to self-test questions in Year 2. Similarly the sighted self-test questions in Year 1 were converted to online quizzes in Year 2 (Figure 1). This was to control for both question difficulty and differences in WebCT question delivery tools.

This study used a formative assessment (online quizzes) and self-tests as a review strategy to enhance student learning. The value of feedback in formative assessment to promote learning rather than just course evaluation is well-known (Blayney & Freeman, 2008; Wiliam & Black, 1996). Feedback in formative assessment exposes knowledge gap or misinformation, which gives students the opportunity to remedy these weaknesses (Wiliam & Black, 1996). It is envisaged that self-test activities would also have the same learning effect, if students were motivated to engage in these voluntary learning activities.

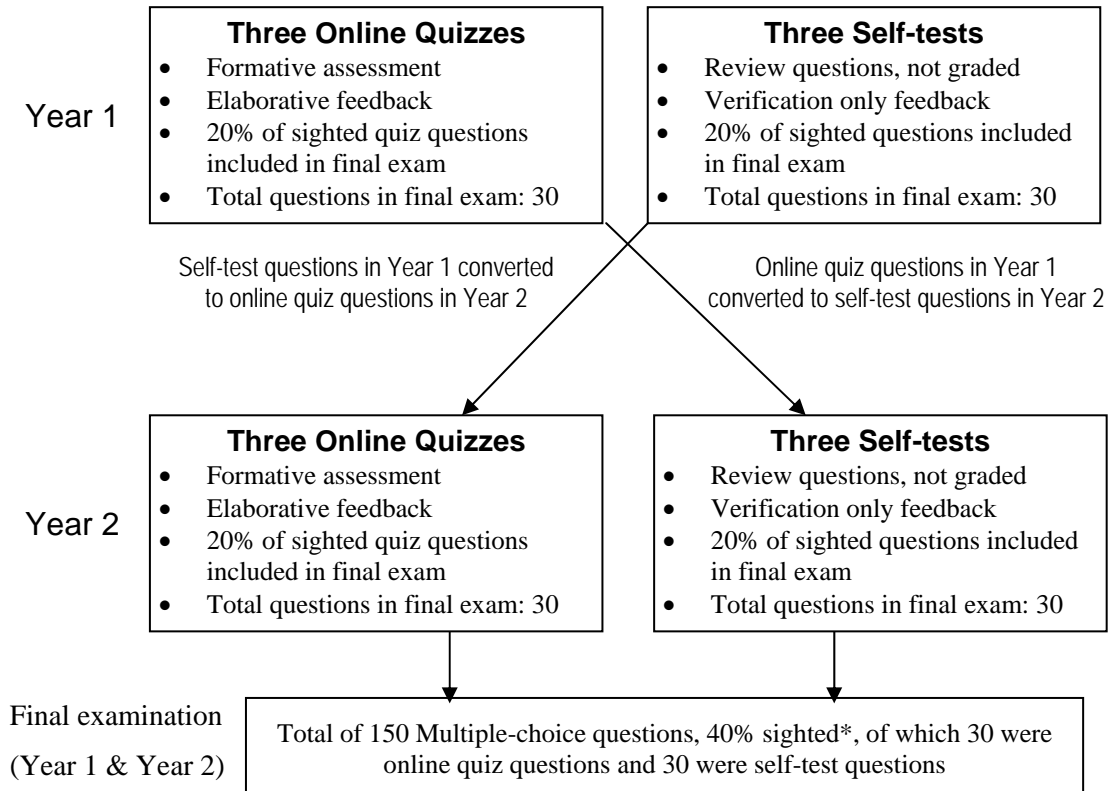
WebCT was the course management platform used to develop the web-based instruction in this course. The course structure of this external web-based elective was divided into three 4-week study periods. At the beginning of each study period, a new content module was uploaded on WebCT. At the same time, a set 50-item self-test quiz was also released, consisting of multiple choice questions with five response options (A, B, C, D, E) each, and this remained available to students for the rest of the semester. Students were encouraged to use this self-test as a tool for revising the module content. The 'answer-until-correct' (AUC) feedback format was used for these self-test questions;

students were required to continue to click until the correct answer was selected, but no other feedback was provided. Students were informed that they could have as many attempts as they wished.

Following the completion of each 4-week study period, an online quiz was time-released for one week, which consisted of 50 multiple-choice questions (with the same five-response format as the self-test) that tested learning content related to that study period. Students were able to complete the online quiz from their home computer or on-campus at their own convenience twice during that week when the quiz was available. The correct answers with elaborative feedback for this formative assessment were made available to the student electronically upon completion and online submission of the quiz. Students could review their attempts as many times as they wished. These were recorded in WebCT as 'online quiz hits'.

There were three online quizzes, each related to the content module of each study period, which contributed to a total weighting of 15% in this elective. Multiple-choice testing has been shown to be an effective measure of student learning (McKeachie, Svinicki, & Hofer, 2006); this study used students' academic performance in the end-of-semester multiple-choice examination as the indicator of the course learning outcome. All formative and summative assessment questions as well as self-tests were of the single-answer, multiple-choice format.

As an added incentive to encourage students to review the self-tests and complete the online quizzes, students were informed at the beginning of the semester that some of the same questions would also be in their final examination, although the percentage of sighted questions to be included in the final examination was not made explicit. Figure 1 presents a summary of the study design.



* The 40% of sighted questions included in the final exam were selected from a total of 437 questions (online quizzes and self-test questions), meaning only 14% of the total pool were included in the final exam.

Figure 1 Study Design

Data analysis

The data were analysed using SPSS for Windows, version 15.0. To assess for differences in student performance in the sighted and unsighted questions as well as the two types of sighted questions, paired *t*-test was used; to compare for differences in exam scores between student characteristics, independent *t*-test and one-way analysis of variance (ANOVA) was used. For comparison of performance in the two types of sighted questions, cumulative GPAs of students were grouped into tertiles. A *P* value of less than 0.05 was considered to indicate statistical significance in this study.

RESULTS

Over the two academic years of the study, 700 students enrolled in this elective. Of these, only 678 (97%) students who completed the end-of-semester final examination were included in the study. There was no significant difference in the mean examination scores between the two years of this study (*P* = 0.53).

Sample characteristics

The mean age of the 678 students was 25.2 years with approximately one-quarter males and three-quarter females (Table 1). Although there were students from 36 different programs of study who undertook this introductory pharmacology elective, nearly two-thirds were nursing students, and approximately one-quarter were enrolled in a health-related program of study. The overall mean examination score was 60.8% and, as expected, the mean score of sighted questions was significantly higher than the mean score of unsighted questions (71.9% versus 51.8%, $P = <0.001$). Subgroup comparison between the two different types of sighted (verification only feedback and elaborative feedback) questions revealed that students performed significantly better in the elaborative feedback questions compared to the verification ('answer-until-correct') only feedback questions (73.7% versus 70.7%, $P = <0.001$).

Characteristic	
Age, (mean \pm SD) years (Range: 18 to 63 years)	25.2 \pm 7.1
Sex, (Male/Female) %	22/78
Program of study clusters ^a %	
• Nursing programs	63
• Applied or health science programs	28
• Other	9
Grade Point Average (mean \pm SD) (Range: 0 to 6.4) ^b	4.1 \pm 1.0
Number of online quiz hits (mean \pm SD) (Range: 0 to 57)	7.4 \pm 7.6
Percentage score in examination (mean \pm SD) (Range: 32 to 95) ^c	60.8 \pm 11.7
Percentage score in unsighted questions (mean \pm SD) (Range: 22 to 96) ^c	51.8 \pm 12.9
Percentage score in sighted questions, (mean \pm SD) (Range: 33 to 100) ^c	71.9 \pm 14.5
• Elaborative feedback questions ^c (mean \pm SD) (Range: 27 to 100)	73.7 \pm 16.3
• Verification only (answer-until-correct) feedback questions ^c (mean \pm SD) (Range: 26 to 100)	70.7 \pm 17.2

^aStudents from 36 different programs of study were enrolled in this elective.

^bPossible GPA range (0 to 7).

^cThese scores were converted to a figure out of 100 to facilitate interpretation.

Table 1 Characteristics of the sample (N = 678)

Engagement in online quizzes

Only 86% of students attempted any of the online quizzes. As expected, those who completed all three online quizzes performed significantly better in their summative examination than those who did not (61.4% versus 57.2%, $P = 0.001$).

Group differences in performance of sighted questions

To compare for differences in examination performance of the sighted questions, the continuous variables that were not normally distributed – age, GPA, and the number of self-test attempts, were dichotomised at the median, as shown in Table 2. Students in the older age group (23 years or older) performed significantly better in both elaborative feedback questions ($P = 0.014$) as well as AUC questions ($P = 0.003$). One-way ANOVA showed that students enrolled in nursing and health science programs performed significantly better than students in other programs, in both ‘elaborative feedback’ as well as AUC sighted questions. In addition, students with higher GPA (>3.25), and students who had more hits on the online quizzes (>5 online quiz hits) performed consistently better in both elaborative feedback as well as AUC sighted questions (Table 2).

Student characteristic	Number of students (n)	Mean scores 'elaborative feedback' questions \pm SD	Mean scores 'answer-until-correct' questions \pm SD
Age			
23 years or older	280	75.6 (\pm 16.3)	73.0 (\pm 17.6)
Less than 23 years old	398	72.4 (\pm 16.1)	69.0 (\pm 16.8)
		$P=0.014$	$P=0.003$
Gender			
Male	146	72.5 (\pm 17.3)	68.5 (\pm 18.4)
Female	532	74.1 (\pm 16.0)	71.3 (\pm 16.8)
		$P=0.319$	$P=0.089$
Type of study program			
Nursing	429	74.9 (\pm 16.3)	72.0 (\pm 17.0)
Applied or health sciences	188	73.3 (\pm 15.0)	70.7 (\pm 16.5)
Other	61	61.3 (\pm 18.4)	61.3 (\pm 18.4)
		$P<0.001$	$P<0.001$
Grade Point Average			
Score > 4.25	328	80.4 (\pm 13.3)	78.5 (\pm 14.8)
Score \leq 4.25	350	67.5 (\pm 16.3)	63.3 (\pm 16.1)
		$P<0.001$	$P<0.001$
Online quiz hits			
> 5 online quiz hits	338	77.1 (\pm 16.0)	78.2 (\pm 14.9)
\leq 5 online quiz hits	340	70.3 (\pm 15.8)	63.2 (\pm 16.2)
		$P<0.001$	$P<0.001$

Table 2 Comparison of students' characteristics with examination performance in 'elaborative feedback' versus 'verification only feedback' questions

Comparison of types of sighted question performance by GPA groupings

Figure 2 presents the differences in students' exam mean scores in the elaborative feedback and verification only feedback questions by GPA groups. As expected, regardless of the type of sighted questions, students with higher GPA tertiles performed consistently better than those in the lower GPA tertiles. Compared to verification only feedback questions, students in the low- and mid-GPA tertiles performed significantly better in the elaborative feedback questions ($P = 0.002$ and $P = 0.001$ respectively). However, this disparity in mean score between the two types of sighted questions was no longer significant among students in the high-GPA tertile ($P = 0.186$).

DISCUSSION

This study was designed to examine the effects of providing response feedback embedded in multiple-choice questions on learning as reflected in students' summative assessment. Results of this study clearly showed the benefits of feedback, a fact that has been well-established since the publication of seminal works like Thorndike (1931) and Skinner (1968). However, what remains uncertain is the varying learning impact of different types of feedback on different student cohorts in different learning contexts. This study shed some light on this issue, as the results demonstrated different levels of effectiveness between the two types of response feedback (verification only and elaborative feedback), depending on student characteristics (William & Black, 1996).

Older students in this study achieved higher scores in both types of sighted questions, which is consistent with previous findings showing the positive relationship between mature-age and better academic performance in higher education (Cantwell, Archer, & Bourke, 2001; Salamonson & Andrew, 2006). The higher mean scores in the sighted questions obtained by students in nursing and health programs were most likely due to prior knowledge of pharmacology. As expected, students with a higher cumulative GPA performed better in both types of sighted questions. Possible explanations for this finding include the inherent scholarly aptitude of these students, and the conscientious study habits of high achievers, which included fully availing themselves of the web-based feedback activities.

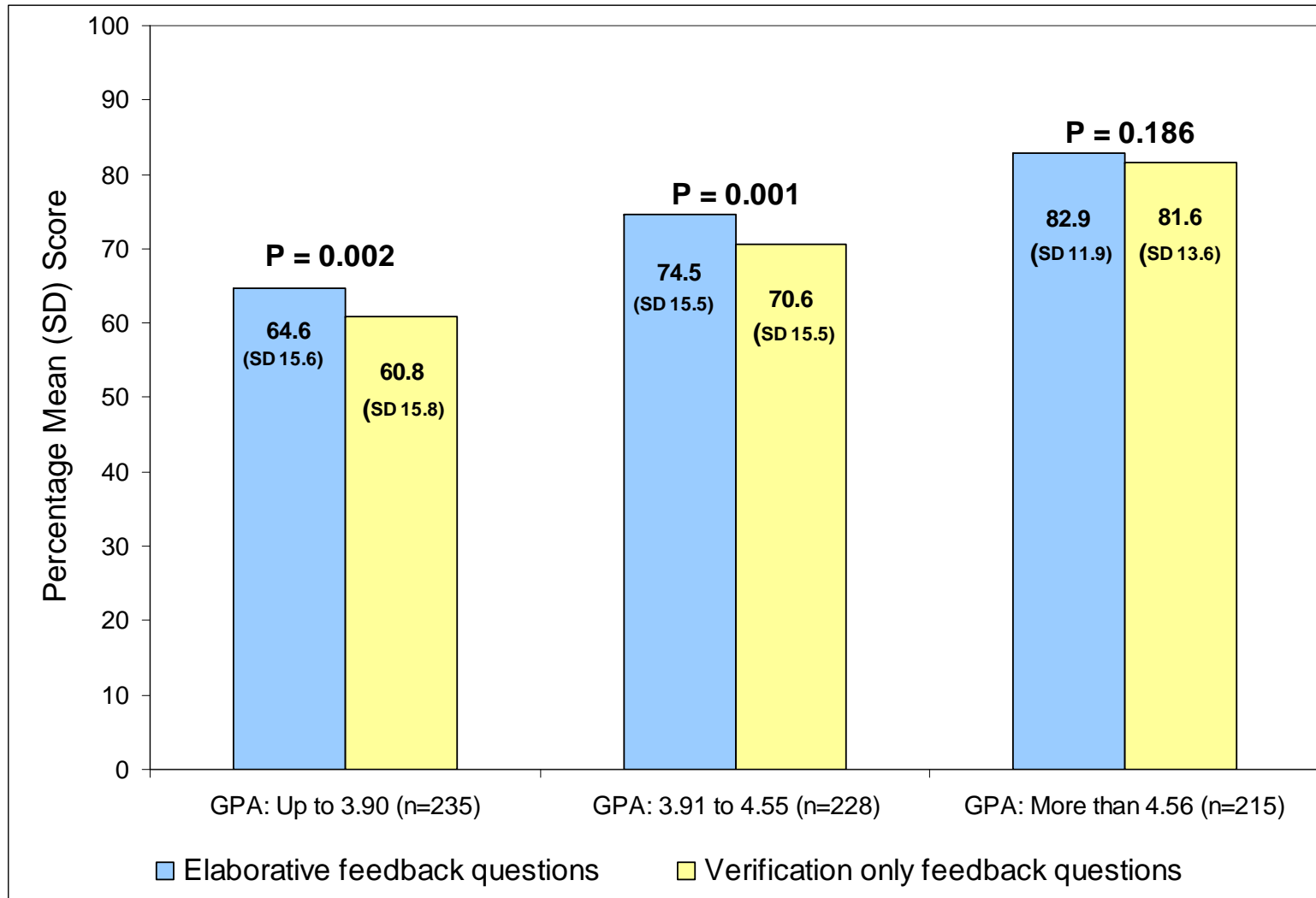


Figure 2 Comparison of performance in 'elaborative feedback' versus 'verification only feedback' questions with GPA groupings

Perhaps the most interesting finding in this study was the difference in mean scores between the two types of sighted questions when compared by GPA tertiles. Even though the trend of higher mean scores in the elaborative feedback questions was consistent across all three GPA groups, this gap was no longer statistically significant in the high tertile group, which gives credence to the assertion by Mason & Bruning (2001) that high achievers need less specific feedback, as they possess the necessary skills to identify errors and are also more likely to actively seek the correct information themselves.

Of concern is the approximate 14% of students who did not complete any of the online quizzes. While it appears that the opportunity for obtaining cumulative marks and the knowledge that some of the same questions would be in the final exam were a sufficient incentives for most students, what remains a challenge is to motivate all students undertaking a web-based course to engage in all feedback learning activities.

The strength of this study was the control of a potential extraneous variable, a difference in question difficulty between the two types of sighted multiple-choice questions. This was controlled by switching the questions used with one type of response feedback in the first academic year to the other type of response format. The limitations of this study include the use of identical questions in the final assessment as this may have only assessed rote learning rather than a deeper understanding of the course material. Further studies examining the effectiveness of feedback, which test the same content with a modification in question phrasing to stimulate deeper processing of learning material, may be useful in elucidating the information process perspective of providing response feedback in multiple-choice questions.

CONCLUSION

This study highlights the effectiveness of providing response feedback in multiple-choice questions, used in formative assessment or self-test activities, particularly elaborative feedback, to promote academic performance in web-based instruction. While the effectiveness of response feedback may depend on student characteristics, those with lower GPAs are likely to derive more benefits from elaborative feedback compared to students with higher GPAs.

REFERENCES

Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61(2), 213-238.

Blayney, P., & Freeman, M. (2008). Individualised interactive formative assessments to promote independent learning. *Journal of Accounting Education, In Press, Corrected Proof*.

Cantwell, R., Archer, J., & Bourke, S. (2001). A comparison of the academic experiences and achievement of university students entering by traditional and non-traditional means. *Assessment and Evaluation in Higher Education, 26*(3), 221-234.

Chang, M. M. (2007). Enhancing web-based language learning through self-monitoring. *Journal of Computer Assisted Learning, 23*(3), 187-196.

Clariana, R., Ross, S., & Morrison, G. (1991). The effects of different feedback strategies using computer-administered multiple-choice questions as instruction. *Educational Technology Research and Development, 39*(2), 5-17.

Cuellar, N. (2002). Tips to increase success for teaching online. *CIN Plus January, 5*(1), 3-6.

Kulhavy, R., & Stock, W. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review, 1*(4), 279-308.

Mandernach, B. J. (2005). Relative effectiveness of computer-based and human feedback for enhancing student learning. *Journal of Educators Online, 2*, 1-12.

Marsh, E. J., Roediger III, H. L., Bjork, R. A., & Bjork, E. L. (2007). The memorial consequences of multiple-choice testing. *Psychonomic Bulletin & Review, 14*(2), 194.

Mason, B. J., & Bruning, R. (2001). *Providing feedback in computer-based instruction: What the research tells us*. Retrieved June 8, 2008, from <http://dwb.unl.edu/Edit/MB/MasonBruning.html>

McKeachie, W. J., Svinicki, M. D., & Hofer, B. K. (2006). *McKeachie's teaching tips: Strategies, research, and theory for college and university teachers*. (12th ed.). Boston, Mass.: Houghton Mifflin Co.

Morrison, G. R., Ross, S. M., Gopalakrishnan, M., & Casey, J. (1995). The effects of feedback and incentives on achievement in computer-based instruction. *Contemporary Educational Psychology, 20*(1), 32-50.

Mory, E. H. (1994). Adaptive feedback in computer-based instruction: Effects of response certitude on performance, feedback-study time, and efficiency. *Journal of Educational Computing Research, 11*(3), 263-290.

Park, O.-C., & Gittelman, S. (1992). Selective use of animation and feedback in computer-based instruction. *Educational Technology Research and Development*, 40(4), 27-38.

Pridemore, D. R., & Klein, J. D. (1995). Control of practice and level of feedback in computer-based instruction. *Contemporary Educational Psychology*, 20(4), 444-450.

Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory. Basic research and implications for educational practice. *Perspectives on Psychological Science*, 1(3), 181-210.

Roper, W. J. (1977). Feedback in computer assisted instruction. *Innovations in Education and Teaching International*, 14(1), 43 - 49.

Salamonson, Y., & Andrew, S. (2006). Academic performance in nursing students: Influence of part-time employment, age and ethnicity. *Journal of Advanced Nursing*, 55(3), 342-349.

Skinner, B. F. (1968). *The technology of teaching*. New York: Appleton-Century-Crofts.

Tallent-Runnels, M. K., Thomas, J. A., Lan, W. Y., Cooper, S., Ahern, T. C., Shaw, S. M., et al. (2006). Teaching courses online: A review of the research. *Review of Educational Research*, 76(1), 93-135.

Thorndike, E. L. (1931). *Human learning*. New York: Appleton-Century-Crofts.

Whyte, M. M., Karolick, D. M., Nielsen, M. C., & Elder, G. D. (1995). Cognitive styles and feedback in computer-assisted instruction. *Journal of Educational Computing Research*, 12(2), 195-203.

William, D., & Black, P. (1996). Meanings and consequences: A basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal*, 22(5), 537-548.